445 Hoes Lane • Piscataway, NJ 08854 USA

IEEE

eXpress Conference Publishing

Premier Publisher of IEEE Conference Proceedings

Sample PDF eBook

for laptop, CD, and USB

About this Publication

Table of Contents

Request an Estimate

October 2012

Editor: D. Graffox

About this Publication

IEEE eXpress Conference Publishing has provided IEEE conferences with an eBook publication option, suitable for both printed and electronic proceedings, for several years. However, once a conference has ended, IEEE conference articles can be obtained electronically only from the IEEE *Xplore*® Digital Library, thereby making it difficult for us to provide an actual eBook product as an example to those unfamiliar with the structure. To remedy this situation, IEEE eXpress has created this mock publication using non-IEEE articles to demonstrate what one may expect from an eBook.

The *technical* articles were collected from several open access publishers and so this publication only models the structure and navigation of the PDF eBook product IEEE eXpress offers and not the scope or content of an IEEE publication. The article content, the format of the articles, and the quality of these PDFs do not necessarily reflect what one might expect of IEEE articles. In fact, all of the original page numbers, running heads, etc., have been retained from the original publisher. Most notably, an IEEE ccc footer (the copyright notice required by the IEEE Letter of Acquisition) cannot appear on these articles. Regardless, these articles serve as stand-ins for technical articles, and IEEE eXpress Conference Publishing makes grateful acknowledgment to the following publishers for their use.

Hindawi Publishing Corporation (see http://www.hindawi.com/journals/)
is the publisher of ***Abstract and Applied Analysis.*** The 48 articles found in Tracks 1 and 2 of this sample eBook are from Volume 2011 beginning at http://www.hindawi.com/journals/aaa/2011/.

Sumy State University (Sumy, Ukraine) (see http://jnep.sumdu.edu.ua/en/component/main/history)
Is the publisher of ***Journal of Nano- and Electronic Physics.*** The 26 articles found in Track 3 of this eBook are from Volume 3, 2011 (http://jnep.sumdu.edu.ua/component/archive/#).

UPGRADE European NETwork (see http://www.cepis.org/upgrade/index.jsp?p=2101&n=2118&a=0)
oversees ***UPGRADE—The European Online Magazine for the IT Professional,*** which provided the Track 4 articles. Specifically, the session 4-1 articles are from "***Privacy in the Net***" (http://www.cepis.org/upgrade/index.jsp?p=0&n=2198&a=3106), the session 4-2 articles are from "***Corporate Websites***" (http://www.cepis.org/upgrade/index.jsp?p=2192&n=2197), the session 4-3 articles are from "***IT and the Transformation of Health Care***" (http://www.cepis.org/upgrade/index.jsp?p=0&n=2189&a=3100), the session 4-4 articles are from "***Visualization of Software***" (http://www.cepis.org/upgrade/index.jsp?p=0&n=2190&a=3101), and the session 4-5 articles are from ***"Voice over IP"*** (http://www.cepis.org/upgrade/index.jsp?p=2146&n=2193).

VSB–Technical University of Ostrava, Faculty of Electrical Engineering and Computer Science, Czech Republic
(see http://advances.utc.sk/index.php/AEEE/about/editorialPolicies#focusAndScope) publishes ***Advances in Electrical and Electronic Engineering.*** The Workshop 1 articles are from Volume 9, Number 5—2011 (http://advances.utc.sk/index.php/AEEE/issue/view/32).

Prime Journals (see http://primejournal.org/about.html)
publishes several on-line journals. The session 6-1 articles are from Volume 2, Issue 1 of ***Prime Research on Education*** (http://primejournal.org/PRE/contents/2012/feb.htm),and the session 6-2 articles are from Volume 1, Issue 12 of ***Prime Journal of Business Administration and Management*** (http://www.primejournal.org/BAM/contents/2011/dec.htm) and Volume 2, Issue 1 (http://www.primejournal.org/BAM/contents/2012/jan.htm).

Contents

Track 1: Abstract and Applied Analysis—Part 1

Session 1-1: Analysis Part 1A

Session 1-2: Analysis Part 1B

to Contents, page 1

to Contents, page 1

Track 2: Abstract and Applied Analysis—Part 2

Session 2-1: Analysis Part 2A

Session 2-2: Analysis Part 2B

Track 3: Nano- and Electronic Physics

Session 3-1: Physics A

Session 3-2: Physics B

Session 3-3: Physics C

Session 3-4: Physics D

Track 4: UPGRADE—The European Online Magazine for the IT Professional

Session 4-1: Privacy in the Net

to Contents, page 1

Session 4-2: Corporate Websites

Session 4-3: IT and the Transformation of Health Care

to Contents, page 1

Session 4-4: Visualization of Software

Session 4-5: Voice over IP

to Contents, page 1

Workshop 1: Advances in Electrical and Electronic Engineering

Workshop 2: Prime Journals

Session 6-1: Prime Research on Education

to Contents, page 1

for Track 2 Listings
for Track 3 Listings
for Track 4 Listings
for Workshop 1 Listings
for Workshop 2 Listings

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 586328, 28 pages
doi:10.1155/2011/586328

Research Article

A Final Result on the Oscillation of Solutions of the Linear Discrete Delayed Equation $\Delta x(n) = -p(n)x(n-k)$ with a Positive Coefficient

J. Baštinec,[1] L. Berezansky,[2] J. Diblík,[1,3] and Z. Šmarda[1]

[1] *Department of Mathematics, Faculty of Electrical Engineering and Communication, Brno University of Technology, 61600 Brno, Czech Republic*
[2] *Department of Mathematics, Ben-Gurion University of the Negev, 84105 Beer-Sheva, Israel*
[3] *Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Brno University of Technology, 602 00 Brno, Czech Republic*

Correspondence should be addressed to J. Diblík, diblik@feec.vutbr.cz

Received 15 January 2011; Accepted 6 June 2011

Academic Editor: Yuri V. Rogovchenko

A linear $(k+1)$th-order discrete delayed equation $\Delta x(n) = -p(n)x(n-k)$ where $p(n)$ a positive sequence is considered for $n \to \infty$. This equation is known to have a positive solution if the sequence $p(n)$ satisfies an inequality. Our aim is to show that, in the case of the opposite inequality for $p(n)$, all solutions of the equation considered are oscillating for $n \to \infty$.

1. Introduction

The existence of positive solutions of difference equations is often encountered when analysing mathematical models describing various processes. This is a motivation for an intensive study of the conditions for the existence of positive solutions of discrete or continuous equations. Such analysis is related to investigating the case of all solutions being oscillating (for investigation in both directions, we refer, e.g., to [1–30] and to the references therein). The existence of monotonous and nontrivial solutions of nonlinear difference equations (the first one implies the existence of solutions of the same sign) also has attracted some attention recently (see, e.g., several, mostly asymptotic methods in [31–42] and the related references therein). In this paper, sharp conditions are derived for all the solutions being oscillating for a class of linear $(k+1)$-order delayed discrete equations.

We consider the delayed $(k+1)$-order linear discrete equation

$$\Delta x(n) = -p(n)x(n-k), \tag{1.1}$$

to Contents, page 1

where $n \in \mathbb{Z}_a^\infty := \{a, a+1, \dots\}$, $a \in \mathbb{N} := \{1, 2, \dots\}$ is fixed, $\Delta x(n) = x(n+1) - x(n)$, $p : \mathbb{Z}_a^\infty \to \mathbb{R}$, $k \in \mathbb{N}$. In what follows, we will also use the sets $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$ and $\mathbb{Z}_a^b := \{a, a+1, \dots, b\}$ for $a, b \in \mathbb{N}$, $a < b$. A solution $x = x(n) : \mathbb{Z}_a^\infty \to \mathbb{R}$ of (1.1) is positive (negative) on $\mathbb{Z}_a^\infty$ if $x(n) > 0$ ($x(n) < 0$) for every $n \in \mathbb{Z}_a^\infty$. A solution $x = x(n) : \mathbb{Z}_a^\infty \to \mathbb{R}$ of (1.1) is oscillating on $\mathbb{Z}_a^\infty$ if it is not positive or negative on $\mathbb{Z}_{a_1}^\infty$ for an arbitrary $a_1 \in \mathbb{Z}_a^\infty$.

Definition 1.1. Let us define the expression $\ln_q t$, $q \geq 1$, by $\ln_q t = \ln(\ln_{q-1} t)$, $\ln_0 t \equiv t$, where $t > \exp_{q-2} 1$ and $\exp_s t = \exp(\exp_{s-1} t)$, $s \geq 1$, $\exp_0 t \equiv t$, and $\exp_{-1} t \equiv 0$ (instead of $\ln_0 t$, $\ln_1 t$, we will only write t and $\ln t$).

In [4] difference equation (1.1) is considered and the following result on the existence of a positive solution is proved.

Theorem 1.2 (see [4]). *Let $q \in \mathbb{N}_0$ be a fixed integer, let $a \in \mathbb{N}$ be sufficiently large and*

$$0 < p(n) \leq \left(\frac{k}{k+1}\right)^k \times \left[\frac{1}{k+1} + \frac{k}{8n^2} + \frac{k}{8(n \ln n)^2} + \cdots + \frac{k}{8(n \ln n \cdots \ln_q n)^2}\right] \tag{1.2}$$

for every $n \in \mathbb{Z}_a^\infty$. Then there exists a positive integer $a_1 \geq a$ and a solution $x = x(n)$, $n \in \mathbb{Z}_{a_1}^\infty$ of equation (1.1) *such that*

$$0 < x(n) \leq \left(\frac{k}{k+1}\right)^n \cdot \sqrt{n \ln n \ln_2 n \cdots \ln_q n} \tag{1.3}$$

holds for every $n \in \mathbb{Z}_{a_1}^\infty$.

Our goal is to answer the open question whether all solutions of (1.1) are oscillating if inequality (1.2) is replaced with the opposite inequality

$$p(n) \geq \left(\frac{k}{k+1}\right)^k \times \left[\frac{1}{k+1} + \frac{k}{8n^2} + \frac{k}{8(n \ln n)^2} + \cdots + \frac{k\theta}{8(n \ln n \cdots \ln_q n)^2}\right] \tag{1.4}$$

assuming $\theta > 1$, and n is sufficiently large. Below we prove that if (1.4) holds and $\theta > 1$, then all solutions of (1.1) are oscillatory. This means that the result given by Theorem 1.2 is a final in a sense. This is discussed in Section 4. Moreover, in Section 3, we show that all solutions of (1.1) will be oscillating if (1.4) holds only on an infinite sequence of subintervals of $\mathbb{Z}_a^\infty$.

Because of its simple form, equation (1.1) (as well as its continuous analogue) attracts permanent attention of investigators. Therefore, in Section 4 we also discuss some of the known results.

The proof of our main result will use the next consequence of one of Domshlak's results [13, Theorem 4, page 66].

Lemma 1.3. *Let s and r be fixed natural numbers such that $r - s > k$. Let $\{\varphi(n)\}_1^\infty$ be a bounded sequence of real numbers and ν_0 be a positive number such that there exists a number $\nu \in (0, \nu_0)$*

to Contents, page 1

satisfying

$$0 \le \sum_{n=s+1}^{i} \varphi(n) \le \frac{\pi}{\nu}, \quad i \in \mathbb{Z}_{s+1}^{r}, \quad \frac{\pi}{\nu} \le \sum_{n=s+1}^{i} \varphi(n) \le \frac{2\pi}{\nu}, \quad i \in \mathbb{Z}_{r+1}^{r+k}, \tag{1.5}$$

$$\varphi(i) \ge 0, \quad i \in \mathbb{Z}_{r+1-k}^{r}, \quad \sum_{n=i+1}^{i+k} \varphi(n) > 0, \quad i \in \mathbb{Z}_{a}^{\infty}, \quad \sum_{n=i}^{i+k} \varphi(n) > 0, \quad i \in \mathbb{Z}_{a}^{\infty}. \tag{1.6}$$

Then, if $p(n) \ge 0$ *for* $n \in \mathbb{Z}_{s+1}^{s+k}$ *and*

$$p(n) \ge \mathcal{R} := \left(\prod_{\ell=n-k}^{n} \frac{\sin\left(\nu \sum_{i=\ell+1}^{\ell+k} \varphi(i)\right)}{\sin\left(\nu \sum_{i=\ell}^{\ell+k} \varphi(i)\right)} \right) \cdot \frac{\sin(\nu\varphi(n-k))}{\sin\left(\nu \sum_{i=n+1-k}^{n} \varphi(i)\right)} \tag{1.7}$$

for $n \in \mathbb{Z}_{s+k+1}^{r}$, *any solution of* (1.1) *has at least one change of sign on* $\mathbb{Z}_{s-k}^{r}$.

Throughout the paper, symbols "o" and "O" (for $n \to \infty$) will denote the well-known Landau order symbols.

2. Main Result

In this section, we give sufficient conditions for all solutions of (1.1) to be oscillatory as $n \to \infty$. It will be necessary to develop asymptotic decompositions of some auxiliary expressions. As the computations needed are rather cumbersome, some auxiliary computations are collected in the appendix to be utilized in the proof of the main result (Theorem 2.1) below.

Theorem 2.1. *Let* $a \in \mathbb{N}$ *be sufficiently large,* $q \in \mathbb{N}_0$ *and* $\theta > 1$. *Assuming that the function* $p : \mathbb{Z}_a^\infty \to (0, \infty)$ *satisfies inequality* (1.4) *for every* $n \in \mathbb{Z}_a^\infty$, *all solutions of* (1.1) *are oscillating as* $n \to \infty$.

Proof. As emphasized above, in the proof, we will use Lemma 1.3. We define

$$\varphi(n) := \frac{1}{n \ln n \ln_2 n \ln_3 n \cdots \ln_q n}, \tag{2.1}$$

where n is sufficiently large, and $q \ge 0$ is a fixed integer. Although the idea of the proof is simple, it is very technical and we will refer to notations and auxiliary computations contained in the appendix. We will develop an asymptotic decomposition of the right-hand side $\mathcal{R}$ of inequality (1.7) with the function $\varphi(n)$ defined by (2.1). We show that this will lead to the desired inequality (1.4). We set

$$\mathcal{R}_1 := \frac{\prod_{i=1}^{k} V(n+i)}{\prod_{i=0}^{k} V^{+}(n+i)} \cdot \varphi(n-k), \tag{2.2}$$

to Contents, page 1

where V and V^+ are defined by (A.4) and (A.5). Moreover, $\mathcal{R}$ can be expressed as

$$\begin{aligned}
\mathcal{R} &= \frac{\sin\left(\nu\sum_{i=n+1-k}^{n}\varphi(i)\right)\prod_{\ell=n-k+1}^{n}\sin\left(\nu\sum_{i=\ell+1}^{\ell+k}\varphi(i)\right)}{\prod_{\ell=n-k}^{n}\sin\left(\nu\sum_{i=\ell}^{\ell+k}\varphi(i)\right)}\cdot\frac{\sin(\nu\varphi(n-k))}{\sin\left(\nu\sum_{i=n+1-k}^{n}\varphi(i)\right)} \\
&= \frac{\prod_{\ell=n-k+1}^{n}\sin\left(\nu\sum_{i=\ell+1}^{\ell+k}\varphi(i)\right)}{\prod_{\ell=n-k}^{n}\sin\left(\nu\sum_{i=\ell}^{\ell+k}\varphi(i)\right)}\cdot\sin(\nu\varphi(n-k)) \\
&= \frac{\prod_{p=1}^{k}\sin(\nu V(n+p))}{\prod_{p=0}^{k}\sin(\nu V^+(n+p))}\cdot\sin(\nu\varphi(n-k)).
\end{aligned} \tag{2.3}$$

Recalling the asymptotic decomposition of $\sin x$ when $x \to 0$: $\sin x = x + O(x^3)$, we get (since $\lim_{n\to\infty}\varphi(n-k) = 0$, $\lim_{n\to\infty}V(n+p) = 0$, $p = 1,\ldots,k$, and $\lim_{n\to\infty}V^+(n+p) = 0$, $p = 0,\ldots,k$)

$$\begin{aligned}
\sin\nu\varphi(n-k) &= \nu\varphi(n-k) + O\left(\nu^3\varphi^3(n-k)\right), \\
\sin\nu V(n+p) &= \nu V(n+p) + O\left(\nu^3 V^3(n+p)\right), \quad p = 1,\ldots,k, \\
\sin\nu V^+(n+p) &= \nu V^+(n+p) + O\left(\nu^3 (V^+)^3(n+p)\right), \quad p = 0,\ldots,k
\end{aligned} \tag{2.4}$$

as $n \to \infty$. Then, it is easy to see that, by (A.13), we have $\varphi(n-\ell) = O(\varphi(n))$, $n \to \infty$ for every $\ell \in \mathbb{R}$ and

$$\mathcal{R} = \mathcal{R}_1\cdot\left(1 + O\left(\nu^2\varphi^2(n)\right)\right), \quad n \longrightarrow \infty. \tag{2.5}$$

Moreover, for $\mathcal{R}_1$, we will get an asymptotic decomposition as $n \to \infty$. Using formulas (A.13), (A.57), and (A.60), we get

$$\begin{aligned}
\mathcal{R}_1 &= \frac{k^k}{(k+1)^{k+1}}\cdot\frac{1-k\alpha(n)-(k/24)(k^2-12k+11)\alpha^2(n)+(k/6)(k^2+5)\sum_{i=0}^{q}\omega_i(n)+O(1/n^3)}{1-(k/24)(k^2+3k+2)\alpha^2(n)+(k/6)(k^2+3k+2)\sum_{i=0}^{q}\omega_i(n)+O(1/n^3)} \\
&\quad\times\left(1 + k\alpha(n) + k^2\sum_{i=0}^{q}\omega_i(n) + O\left(\frac{1}{n^3}\right)\right).
\end{aligned} \tag{2.6}$$

Since $\lim_{n\to\infty}\alpha(n) = 0$, $\lim_{n\to\infty}\omega_i(n) = 0$, $i = 1,\ldots,q$, we can decompose the denominator of the second fraction as the sum of the terms of a geometric sequence. Keeping only terms with

to Contents, page 1

the order of accuracy necessary for further analysis (i.e. with order $O(1/n^3)$), we get

$$\begin{aligned}
&\left(1-\frac{k}{24}\left(k^2+3k+2\right)\alpha^2(n)+\frac{k}{6}\left(k^2+3k+2\right)\sum_{i=0}^{q}\omega_i(n)+O\left(\frac{1}{n^3}\right)\right)^{-1}\\
&\quad=1+\frac{k}{24}\left(k^2+3k+2\right)\alpha^2(n)-\frac{k}{6}\left(k^2+3k+2\right)\sum_{i=0}^{q}\omega_i(n)+O\left(\frac{1}{n^3}\right).
\end{aligned}\tag{2.7}$$

We perform an auxiliary computation in $\mathcal{R}_1$,

$$\begin{aligned}
&\left(1-k\alpha(n)-\frac{k}{24}\left(k^2-12k+11\right)\alpha^2(n)+\frac{k}{6}\left(k^2+5\right)\sum_{i=0}^{q}\omega_i(n)+O\left(\frac{1}{n^3}\right)\right)\\
&\quad\times\left(1+\frac{k}{24}\left(k^2+3k+2\right)\alpha^2(n)-\frac{k}{6}\left(k^2+3k+2\right)\sum_{i=0}^{q}\omega_i(n)+O\left(\frac{1}{n^3}\right)\right)\\
&\quad\times\left(1+k\alpha(n)+k^2\sum_{i=0}^{q}\omega_i(n)+O\left(\frac{1}{n^3}\right)\right)\\
&\quad=\left(1-k\alpha(n)-\frac{k}{24}\left(k^2-12k+11\right)\alpha^2(n)+\frac{k}{6}\left(k^2+5\right)\sum_{i=0}^{q}\omega_i(n)+O\left(\frac{1}{n^3}\right)\right)\\
&\qquad\times\left(1+k\alpha(n)+\frac{k}{24}\left(k^2+3k+2\right)\alpha^2(n)-\frac{k}{6}\left(k^2-3k+2\right)\sum_{i=0}^{q}\omega_i(n)+O\left(\frac{1}{n^3}\right)\right)\\
&\quad=1-\frac{3}{8}k(k+1)\alpha^2(n)+\frac{1}{2}k(k+1)\sum_{i=0}^{q}\omega_i(n)+O\left(\frac{1}{n^3}\right)=(\text{we use formula (A.15)})\\
&\quad=1+\frac{1}{8}k(k+1)\Omega(n)+O\left(\frac{1}{n^3}\right)\\
&\quad=1+\frac{1}{8}k(k+1)\left(\frac{1}{n^2}+\frac{1}{(n\ln n)^2}+\frac{1}{(n\ln n\ln_2 n)^2}+\cdots+\frac{1}{(n\ln n\ln_2 n\cdots\ln_q n)^2}\right)\\
&\qquad+O\left(\frac{1}{n^3}\right).
\end{aligned}\tag{2.8}$$

Thus, we have

$$\begin{aligned}
\mathcal{R}_1&=\frac{k^k}{(k+1)^{k+1}}\times\left[1+\frac{1}{8}k(k+1)\left(\frac{1}{n^2}+\frac{1}{(n\ln n)^2}+\frac{1}{(n\ln n\ln_2 n)^2}+\cdots\right.\right.\\
&\qquad\left.\left.+\frac{1}{(n\ln n\ln_2 n\cdots\ln_q n)^2}\right)\right]+O\left(\frac{1}{n^3}\right)\\
&=\left(\frac{k}{k+1}\right)^k\times\left[\frac{1}{k+1}+\frac{k}{8n^2}+\frac{k}{8(n\ln n)^2}+\cdots+\frac{k}{8(n\ln n\cdots\ln_q n)^2}\right]+O\left(\frac{1}{n^3}\right).
\end{aligned}\tag{2.9}$$

to Contents, page 1

Finalizing our decompositions, we see that

$$
\begin{aligned}
\mathcal{R} &= \mathcal{R}_1 \cdot \left(1 + O\left(\nu^2\varphi^2(n)\right)\right) \\
&= \left(\left(\frac{k}{k+1}\right)^k \times \left[\frac{1}{k+1} + \frac{k}{8n^2} + \frac{k}{8(n\ln n)^2} + \cdots + \frac{k}{8(n\ln n\cdots\ln_q n)^2}\right] + O\left(\frac{1}{n^3}\right)\right) \\
&\quad \times \left(1 + O\left(\nu^2\varphi^2(n)\right)\right) \\
&= \left(\frac{k}{k+1}\right)^k \times \left[\frac{1}{k+1} + \frac{k}{8n^2} + \frac{k}{8(n\ln n)^2} + \cdots + \frac{k}{8(n\ln n\cdots\ln_q n)^2}\right] \\
&\quad + O\left(\frac{\nu^2}{(n\ln n\cdots\ln_q n)^2}\right).
\end{aligned}
\tag{2.10}
$$

It is easy to see that inequality (1.7) becomes

$$
\begin{aligned}
p(n) \geq \left(\frac{k}{k+1}\right)^k \times \left[\frac{1}{k+1} + \frac{k}{8n^2} + \frac{k}{8(n\ln n)^2} + \cdots + \frac{k}{8(n\ln n\cdots\ln_q n)^2}\right] \\
+ O\left(\frac{\nu^2}{(n\ln n\cdots\ln_q n)^2}\right)
\end{aligned}
\tag{2.11}
$$

and will be valid if (see (1.4))

$$
\begin{aligned}
&\frac{1}{k+1} + \frac{k}{8n^2} + \frac{k}{8(n\ln n)^2} + \cdots + \frac{k\theta}{8(n\ln n\cdots\ln_q n)^2} \\
&\quad \geq \frac{1}{k+1} + \frac{k}{8n^2} + \frac{k}{8(n\ln n)^2} + \cdots + \frac{k}{8(n\ln n\cdots\ln_q n)^2} + O\left(\frac{\nu^2}{(n\ln n\cdots\ln_q n)^2}\right)
\end{aligned}
\tag{2.12}
$$

or

$$
\theta \geq 1 + O\left(\nu^2\right)
\tag{2.13}
$$

for $n \to \infty$. If $n \geq n_0$, where n_0 is sufficiently large, (2.13) holds for $\nu \in (0, \nu_0)$ with ν_0 sufficiently small because $\theta > 1$. Consequently, (2.11) is satisfied and the assumption (1.7) of Lemma 1.3 holds for $n \in \mathbb{Z}_{n_0}^{\infty}$. Let $s \geq n_0$ in Lemma 1.3 be fixed, $r > s + k + 1$ be so large (and ν_0 so small if necessary) that inequalities (1.5) hold. Such choice is always possible since the series $\sum_{n=s+1}^{\infty} \varphi(n)$ is divergent. Then Lemma 1.3 holds and any solution of (1.1) has at least one change of sign on $\mathbb{Z}_{s-k}^{r}$. Obviously, inequalities (1.5) can be satisfied for another pair of (s, r), say (s_1, r_1) with $s_1 > r$ and $r_1 > s_1 + k$ sufficiently large and, by Lemma 1.3, any solution of (1.1) has at least one change of sign on $\mathbb{Z}_{s_1-k}^{r_1}$. Continuing this process, we will get a sequence of pairs (s_j, r_j) with $\lim_{j\to\infty} s_j = \infty$ such that any solution of (1.1) has at least one change of sign on $\mathbb{Z}_{s_j-k}^{r_j}$. This concludes the proof. □

to Contents, page 1

3. A Generalization

The coefficient p in Theorem 2.1 is supposed to be positive on $\mathbb{Z}_a^\infty$. For all sufficiently large n, the expression $\mathcal{R}$, as can easily be seen from (2.10), is positive. Then, as follows from Lemma 1.3, any solution of (1.1) has at least one change of sign on $\mathbb{Z}_{s-k}^r$ if p is nonnegative on $\mathbb{Z}_{s+1}^{s+k}$ and satisfies inequality (1.4) on $\mathbb{Z}_{s+k+1}^r$.

Owing to this remark, Theorem 2.1 can be generalized because (and the following argumentation was used at the end of the proof of Theorem 2.1) all solutions of (1.1) will be oscillating as $n \to \infty$ if a sequence of numbers $\{s_i, r_i\}$, $r_i > s_i + k + 1$, $s_1 \geq a$, $i = 1, 2, \dots$ exists such that $s_{i+1} > r_i$ (i.e., the sets $\mathbb{Z}_{s_i}^{r_i}$, $\mathbb{Z}_{s_{i+1}}^{r_{i+1}}$ are disjoint and $\lim_{i\to\infty} s_i = \infty$), and, for every pair (s_i, r_i), all assumptions of Lemma 1.3 are satisfied (because of the specification of function φ by (2.1), inequalities (1.6) are obviously satisfied). This means that, on the set

$$\mathcal{M} := \mathbb{Z}_a^\infty \setminus \bigcup_{i=1}^{\infty} \mathbb{Z}_{s_i}^{r_i}, \tag{3.1}$$

function p can assume even negative values, and, moreover, there is no restriction on the behavior of $p(n)$ for $n \in \mathcal{M}$. This leads to the following generalization of Theorem 2.1 with a proof similar to that of Theorem 2.1 and, therefore, omitted.

Theorem 3.1. *Let $a \in \mathbb{N}$ be sufficiently large, $q \in \mathbb{N}_0$, ν_0 be a positive number, $\theta > 1$ and $p : \mathbb{Z}_a^\infty \to \mathbb{R}$. Let there exists a sequence on integers $\{s_j, r_j\}$, $r_j > s_j + k + 1$, $j = 1, 2, \dots$, $s_1 \geq a$, s_1 sufficiently large and $s_{j+1} > r_j$ such that, for function φ (defined by (2.1)) and for each pair (s_j, r_j), $j = 1, 2, \dots$, there exists a number $\nu_j \in (0, \nu_0)$ such that*

$$0 \leq \sum_{n=s_j+1}^{i} \varphi(n) \leq \frac{\pi}{\nu_j}, \quad i \in \mathbb{Z}_{s_j+1}^{r_j}, \quad \frac{\pi}{\nu_j} \leq \sum_{n=s_j+1}^{i} \varphi(n) \leq \frac{2\pi}{\nu_j}, \quad i \in \mathbb{Z}_{r_j+1}^{r_j+k}, \tag{3.2}$$

$p(n) \geq 0$ for $n \in \mathbb{Z}_{s_j+1}^{s_j+k}$, and (1.4) holds for $n \in \mathbb{Z}_{s_j+k+1}^{r_j}$, then all solutions of (1.1) are oscillating as $n \to \infty$.

4. Comparisons, Concluding Remarks, and Open Problems

Equation (1.1) with $k = 1$ was considered in [5], where a particular case of Theorem 2.1 is proved. In [4], a hypothesis is formulated together with the proof of Theorem 1.2 (Conjecture 1) about the oscillation of all solutions of (1.1) almost coinciding with the formulation of Theorem 2.1. For its simple form, (1.1) is often used for testing new results and is very frequently investigated.

Theorems 1.2 and 2.1 obviously generalize several classical results. We mention at least some of the simplest ones (see, e.g., [16, Theorem 7.7] or [19, Theorem 7.5.1]),

Theorem 4.1. *Let $p(n) \equiv p = \text{const}$. Then every solution of (1.1) oscillates if and only if*

$$p > \frac{k^k}{(k+1)^{k+1}}. \tag{4.1}$$

Or the following result holds as well (see, e.g., [16, Theorem 7.6]) [18, 19]).

to Contents, page 1

Theorem 4.2. *Let* $p(n) \geq 0$ *and*

$$\sup_n p(n) < \frac{k^k}{(k+1)^{k+1}}. \tag{4.2}$$

Then (1.1) *has a nonoscillatory solution.*

In [9] a problem on oscillation of all solutions of equation

$$\Delta u(n) + p(n)u(\tau(n)) = 0, \quad n \in \mathbb{N} \tag{4.3}$$

is considered where $p : \mathbb{N} \to \mathbb{R}_+$, $\tau : \mathbb{N} \to \mathbb{N}$, and $\lim_{n\to\infty}\tau(n) = +\infty$. Since in (4.3) delay τ is variable, we can formulate

Open Problem 1. It is an interesting open question whether Theorems 1.2 and 2.1 can be extended to linear difference equations with a variable delay argument of the form, for example,

$$\Delta u(n) = -p(n)u(h(n)), \quad n \in \mathbb{Z}_a^\infty, \tag{4.4}$$

where $0 \leq n - h(n) \leq k$. For some of the related results for the differential equation

$$\dot{x}(t) = -p(t)x(h(t)), \tag{4.5}$$

see the results in [3, 12] that are described below.

Open Problem 2. It is well known [19, 22] that the following condition is also sufficient for the oscillation of all solutions of (4.5) with $h(n) = n - k$:

$$\liminf_{n\to\infty} \frac{1}{k} \sum_{i=n-k}^{n-1} p_i > \frac{k^k}{(k+1)^{k+1}}. \tag{4.6}$$

The right-hand side of (4.6) is a critical value for this criterion since this number cannot be replaced with a smaller one.

In [30] equation (1.1) is considered as well. The authors prove that all solutions oscillate if $p(n) \geq 0$, $\varepsilon > 0$ and

$$\limsup_{n\to\infty} p(n) > \frac{k^k}{(k+1)^{k+1}} - \frac{\varepsilon}{k} + 4k\varepsilon^{1/4}, \tag{4.7}$$

where

$$\varepsilon = \left(\frac{k}{k+1}\right)^{k+1} - \liminf_{n\to\infty} \sum_{i=n-k}^{n-1} p_i. \tag{4.8}$$

to Contents, page 1

An open problem is to obtain conditions similar to Theorem 2.1 for this kind of oscillation criteria. Some results on this problem for delay differential equations were also obtained in paper [3].

In [26] the authors establish an equivalence between the oscillation of (1.1) and the equation

$$\Delta^2 y(n-1) + \frac{2(k+1)^k}{k^{k+1}}\left(p(n) - \frac{k^k}{(k+1)^{k+1}}\right)y(n) = 0 \tag{4.9}$$

under the critical state

$$\liminf_{n\to\infty} p(n) = \frac{k^k}{(k+1)^{k+1}}, \tag{4.10}$$

$$p(n) \geq \frac{k^k}{(k+1)^{k+1}}. \tag{4.11}$$

Then they obtain some sharp oscillation and nonoscillation criteria for (1.1). One of the results obtained there is the following.

Theorem 4.3. *Assume that, for sufficiently large n, inequality* (4.11) *holds. Then the following statements are valid.*

(i) *If*

$$\liminf_{n\to\infty}\left[\left(p(n) - \frac{k^k}{(k+1)^{k+1}}\right)n^2\right] > \frac{k^{k+1}}{8(k+1)^k}, \tag{4.12}$$

then every solution of (1.1) *is oscillatory.*

(ii) *If, on the other hand,*

$$\left(p(n) - \frac{k^k}{(k+1)^{k+1}}\right)n^2 \leq \frac{k^{k+1}}{8(k+1)^k}, \tag{4.13}$$

then (1.1) *has a nonoscillatory solution.*

Regarding our results, it is easy to see that statement (i) is a particular case of Theorem 2.1 while statement (ii) is a particular case of Theorem 1.2.

In [27], the authors investigate (1.1) for $n \geq n_0$ and prove that (1.1) is oscillatory if

$$\sum_{i=n_0}^{\infty} p(i)\left\{\frac{k+1}{k}\cdot\sqrt[k+1]{\sum_{j=i+1}^{i+k} p(j)} - 1\right\} = \infty. \tag{4.14}$$

to Contents, page 1

Comparing (4.14) with Theorem 2.1, we can see that (4.14) gives not sharp sufficient condition. Set, for example, $k = 1$, $\theta > 1$ and

$$p(n) = \frac{1}{2}\left[\frac{1}{2} + \frac{\theta}{8n^2}\right]. \tag{4.15}$$

Then,

$$\frac{k+1}{k} \cdot \sqrt[k+1]{\sum_{j=i+1}^{i+k} p(j)} - 1 = 2 \cdot \sqrt{\frac{1}{4}\left(1 + \frac{\theta}{4(i+1)^2}\right)} - 1 = \frac{\theta/4(i+1)^2}{1 + \sqrt{1 + \theta/4(i+1)^2}} \tag{4.16}$$

and the series in the left-hand side of (4.14) converges since

$$\begin{aligned}
&\sum_{i=n_0}^{\infty} p(i)\left\{\frac{k+1}{k} \sqrt[k+1]{\sum_{j=i+1}^{i+k} p(j)} - 1\right\} \\
&\quad = \sum_{i=n_0}^{\infty} \frac{1}{2}\left[\frac{1}{2} + \frac{\theta}{8i^2}\right] \frac{\theta/4(i+1)^2}{1 + \sqrt{1 + \theta/4(i+1)^2}} \le \theta \sum_{i=n_0}^{\infty} \frac{1}{i^2}\left[1 + \frac{\theta}{i^2}\right] < \infty.
\end{aligned} \tag{4.17}$$

But, by Theorem 2.1 all solutions of (1.1) are oscillating as $n \to \infty$. Nevertheless (4.14) is not a consequence of Theorem 2.1.

Let us consider a continuous variant of (1.1): a delayed differential linear equation of the form

$$\dot{x}(t) = -a(t)x(t-\tau), \tag{4.18}$$

where $\tau > 0$ is a constant delay and $a : [t_0, \infty) \to (0, \infty)$ (or $a : [t_0, \infty) \to \mathbb{R}$), $t_0 \in \mathbb{R}$. This equation, too, for its simple form, is often used for testing new results and is very frequently investigated. It is, for example, well known that a scalar linear equation with delay

$$\dot{x}(t) + \frac{1}{e}x(t-1) = 0 \tag{4.19}$$

has a nonoscillatory solution as $t \to \infty$. This means that there exists an eventually positive solution. The coefficient $1/e$ is called critical with the following meaning: for any $\alpha > (1/e)$, all solutions of the equation

$$\dot{x}(t) + \alpha x(t-1) = 0 \tag{4.20}$$

are oscillatory while, for $\alpha \le (1/e)$, there exists an eventually positive solution. In [10], the third author considered (4.18), where $a : [t_0, \infty) \to (0, \infty)$ is a continuous function, and t_0 is sufficiently large. For the critical case, he obtained the following result (being a continuous variant of Theorems 1.2 and 2.1).

to Contents, page 1

Theorem 4.4. *(a) Let an integer $k \geq 0$ exist such that $a(t) \leq a_k(t)$ if $t \longrightarrow \infty$ where*

$$a_k(t) := \frac{1}{e\tau} + \frac{\tau}{8et^2} + \frac{\tau}{8e(t\ln t)^2} + \cdots + \frac{\tau}{8e(t\ln t\ln_2 t\cdots\ln_k t)^2}. \tag{4.21}$$

Then there exists an eventually positive solution x of (4.18).

(b) Let an integer $k \geq 2$ and $\theta > 1$, $\theta \in \mathbb{R}$ exist such that

$$a(t) > a_{k-2}(t) + \frac{\theta\tau}{8e(t\ln t\ln_2 t\cdots\ln_{k-1} t)^2}, \tag{4.22}$$

if $t \longrightarrow \infty$. Then all solutions of (4.18) *oscillate.*

Further results on the critical case for (4.18) can be found in [1, 11, 14, 17, 24].

In [12], Theorem 7 was generalized for equations with a variable delay

$$\dot{x}(t) + a(t)x(t - \tau(t)) = 0, \tag{4.23}$$

where $a : [t_0, \infty) \longrightarrow (0, \infty)$ and $\tau : [t_0, \infty) \longrightarrow (0, \infty)$ are continuous functions. The main results of this paper include the following.

Theorem 4.5 (see [12]). *Let $t - \tau(t) \geq t_0 - \tau(t_0)$ if $t \geq t_0$. Let an integer $k \geq 0$ exist such that $a(t) \leq a_{k\tau}(t)$ for $t \longrightarrow \infty$, where*

$$a_{k\tau}(t) := \frac{1}{e\tau(t)} + \frac{\tau(t)}{8et^2} + \frac{\tau(t)}{8e(t\ln t)^2} + \cdots + \frac{\tau(t)}{8e(t\ln t\ln_2 t\cdots\ln_k t)^2}. \tag{4.24}$$

If moreover

$$\begin{gathered} \int_{t-\tau(t)}^{t} \frac{1}{\tau(\xi)} d\xi \leq 1 \quad \text{when } t \longrightarrow \infty, \\ \lim_{t\to\infty} \tau(t) \cdot \left(\frac{1}{t}\ln t\ln_2 t\cdots\ln_k t\right) = 0, \end{gathered} \tag{4.25}$$

then there exists an eventually positive solution x of (4.23) *for $t \longrightarrow \infty$.*

Finally, the last results were generalized in [3]. We reproduce some of the results given there.

Theorem 4.6. *(A) Let $\tau > 0$, $0 \leq \tau(t) \leq \tau$ for $t \longrightarrow \infty$, and let condition (a) of Theorem 4.4 holds. Then* (4.23) *has a nonoscillatory solution.*

(B) Let $\tau(t) \geq \tau > 0$ for $t \longrightarrow \infty$, and let condition (b) of Theorem 4.4 holds. Then all solutions of (4.23) *oscillate.*

For every integer $k \geq 0$, $\delta > 0$ and $t \longrightarrow \infty$, we define

$$A_k(t) := \frac{1}{e\delta\tau(t)} + \frac{\delta}{8e\tau(t)s^2} + \frac{\delta}{8e\tau(t)(s\ln s)^2} + \cdots + \frac{\delta}{8e\tau(t)(s\ln s\ln_2 s\cdots\ln_k s)^2}, \tag{4.26}$$

to Contents, page 1

where

$$s = p(t) := \int_{t_0}^{t} \frac{1}{\tau(\xi)} d\xi. \tag{4.27}$$

Theorem 4.7. *Let for t_0 sufficiently large and $t \geq t_0$: $\tau(t) > 0$ a.e., $1/\tau(t)$ be a locally integrable function,*

$$\lim_{t\to\infty} (t - \tau(t)) = \infty, \quad \int_{t_0}^{\infty} \frac{1}{\tau(\xi)} d\xi = \infty, \tag{4.28}$$

and let there exists $t_1 > t_0$ such that $t - \tau(t) \geq t_0$, $t \geq t_1$.

(a) *If there exists a $\delta \in (0, \infty)$ such that*

$$\int_{t-\tau(t)}^{t} \frac{1}{\tau(\xi)} d\xi \leq \delta, \quad t \geq t_1, \tag{4.29}$$

and, for a fixed integer $k \geq 0$,

$$a(t) \leq A_k(t), \quad t \geq t_1, \tag{4.30}$$

then there exists an eventually positive solution of (4.23)*.*

(b) *If there exists a $\delta \in (0, \infty)$ such that*

$$\int_{t-\tau(t)}^{t} \frac{1}{\tau(\xi)} d\xi \geq \delta, \quad t \geq t_1, \tag{4.31}$$

and, for a fixed integer $k \geq 2$ and $\theta > 1$, $\theta \in \mathbb{R}$,

$$a(t) > A_{k-2}(t) + \frac{\theta\delta}{8e\tau(t)(s \ln s \ln_2 s \cdots \ln_{k-1} s)^2}, \tag{4.32}$$

if $t \geq t_1$, then all solutions of (4.23) *oscillate.*

Appendix

A. Auxiliary Computations

This part includes auxiliary results with several technical lemmas proved. Part of them is related to the asymptotic decomposition of certain functions and the rest deals with computing the sums of some algebraic expressions. The computations are referred to in the proof of the main result (Theorem 2.1) in Section 2.

to Contents, page 1

First we define auxiliary functions (recalling also the definition of function φ given by (2.1)):

$$\begin{aligned}
\varphi(n) &:= \frac{1}{n \ln n \ln_2 n \ln_3 n \cdots \ln_q n}, \\
\alpha(n) &:= \frac{1}{n} + \frac{1}{n \ln n} + \frac{1}{n \ln n \ln_2 n} + \cdots + \frac{1}{n \ln n \ln_2 n \cdots \ln_q n}, \\
\omega_0(n) &:= \frac{1}{n^2} + \frac{3}{2n^2 \ln n} + \frac{3}{2n^2 \ln n \ln_2 n} + \cdots + \frac{3}{2n^2 \ln n \ln_2 n \cdots \ln_q n}, \\
\omega_1(n) &:= \frac{1}{(n \ln n)^2} + \frac{3}{2(n \ln n)^2 \ln_2 n} + \cdots + \frac{3}{2(n \ln n)^2 \ln_2 n \cdots \ln_q n}, \\
&\vdots \\
\omega_{q-1}(n) &:= \frac{1}{(n \ln n \cdots \ln_{q-1} n)^2} + \frac{3}{2(n \ln n \cdots \ln_{q-1} n)^2 \ln_q n}, \\
\omega_q(n) &:= \frac{1}{(n \ln n \cdots \ln_q n)^2}, \\
\Omega(n) &:= \frac{1}{n^2} + \frac{1}{(n \ln n)^2} + \frac{1}{(n \ln n \ln_2 n)^2} + \cdots + \frac{1}{(n \ln n \ln_2 n \cdots \ln_q n)^2},
\end{aligned} \tag{A.1}$$

where n is sufficiently large and $q \in \mathbb{N}_0$. Moreover, we set (for admissible values of arguments)

$$\Sigma(p) := \sum_{\ell=1}^{k} (k - p - \ell), \tag{A.2}$$

$$\Sigma^+(p) := \Sigma(p) + (k - p), \tag{A.3}$$

$$V(n+p) := \sum_{\ell=1}^{k} \varphi(n + p - k + \ell), \tag{A.4}$$

$$V^+(n+p) := V(n+p) + \varphi(n + p - k), \tag{A.5}$$

$$S(p) := \sum_{\ell=1}^{k} (k - p - \ell)^2, \tag{A.6}$$

$$S^+(p) := S(p) + (k - p)^2. \tag{A.7}$$

A.1. Asymptotic Decomposition of Iterative Logarithms

In the proof of the main result, we use auxiliary results giving asymptotic decompositions of iterative logarithms. The following lemma is proved in [11].

to Contents, page 1

Lemma A.1. *For fixed* $r, \sigma \in \mathbb{R} \setminus \{0\}$ *and a fixed integer* $s \geq 1$, *the asymptotic representation*

$$\begin{aligned}\frac{\ln_s^{\sigma}(n-r)}{\ln_s^{\sigma} n} = 1 &- \frac{r\sigma}{n \ln n \cdots \ln_s n} - \frac{r^2\sigma}{2n^2 \ln n \cdots \ln_s n} \\ &- \frac{r^2\sigma}{2(n\ln n)^2 \ln_2 n \cdots \ln_s n} - \cdots - \frac{r^2\sigma}{2(n \ln n \cdots \ln_{s-1} n)^2 \ln_s n} \\ &+ \frac{r^2\sigma(\sigma-1)}{2(n\ln n \cdots \ln_s n)^2} - \frac{r^3\sigma(1+o(1))}{3n^3 \ln n \cdots \ln_s n}\end{aligned} \tag{A.8}$$

holds for $n \to \infty$.

A.2. Formulas for $\Sigma(p)$ and for $\Sigma^+(p)$

Lemma A.2. *The following formulas hold:*

$$\Sigma(p) = \frac{k}{2} \cdot (k-2p-1), \tag{A.9}$$

$$\Sigma^+(p) = \frac{k+1}{2} \cdot (k-2p). \tag{A.10}$$

Proof. It is easy to see that

$$\begin{aligned}\Sigma(p) &= \sum_{\ell=-p}^{k-p-1} \ell = (k-p-1) + (k-p-2) + \cdots + (-p) \\ &= (k-(p+1)) + (k-(p+2)) + \cdots + (k-(p+k)) = \frac{k}{2} \cdot (k-2p-1), \\ \Sigma^+(p) &= \Sigma(p) + (k-p) = \frac{k+1}{2} \cdot (k-2p).\end{aligned} \tag{A.11}$$

□

A.3. Formula for the Sum of the Terms of an Arithmetical Sequence

Denote by $u_1, u_2, \ldots, u_r$ the terms of an arithmetical sequence of kth order (kth differences are constant), $d_1', d_2', d_3', \ldots$, the first differences ($d_1' = u_2 - u_1$, $d_2' = u_3 - u_2, \ldots$), $d_1'', d_2'', d_3'', \ldots$, the second differences ($d_1'' = d_2' - d_1', \ldots$), and so forth. Then the following result holds (see, e.g., [43]).

Lemma A.3. *For the sum of* r *terms of an arithmetical sequence of* k*th order, the following formula holds*

$$\sum_{i=1}^{r} u_i = \frac{r!}{(r-1)! \cdot 1!} \cdot u_1 + \frac{r!}{(r-2)! \cdot 2!} \cdot d_1' + \frac{r!}{(r-3)! \cdot 3!} \cdot d_1'' + \cdots. \tag{A.12}$$

to Contents, page 1

A.4. Asymptotic Decomposition of $\varphi(n-l)$

Lemma A.4. *For fixed* $\ell \in \mathbb{R}$ *and* $q \in \mathbb{N}_0$*, the asymptotic representation*

$$\varphi(n-\ell) = \varphi(n)\left(1 + \ell\alpha(n) + \ell^2 \sum_{i=0}^{q} \omega_q(n)\right) + O\left(\frac{\varphi(n)}{n^3}\right) \tag{A.13}$$

holds for $n \to \infty$.

Proof. The function $\varphi(n)$ is defined by (2.1). We develop the asymptotic decomposition of $\varphi(n-\ell)$ when n is sufficiently large and $\ell \in \mathbb{R}$. Applying Lemma A.1 (for $\sigma = -1$, $r = \ell$ and $s = 1, 2, \ldots, q$), we get

$$\begin{aligned}
\varphi(n-\ell) &= \frac{1}{(n-\ell)\ln(n-\ell)\ln_2(n-\ell)\ln_3(n-\ell)\cdots\ln_q(n-\ell)} \\
&= \frac{1}{n(1-\ell/n)\ln(n-\ell)\ln_2(n-\ell)\ln_3(n-\ell)\cdots\ln_q(n-\ell)} \\
&= \varphi(n)\cdot\frac{1}{1-\ell/n}\cdot\frac{\ln n}{\ln(n-\ell)}\cdot\frac{\ln_2 n}{\ln_2(n-\ell)}\cdot\frac{\ln_3 n}{\ln_3(n-\ell)}\cdots\frac{\ln_q n}{\ln_q(n-\ell)} \\
&= \varphi(n)\left(1 + \frac{\ell}{n} + \frac{\ell^2}{n^2} + O\left(\frac{1}{n^3}\right)\right) \\
&\quad\times\left(1 + \frac{\ell}{n\ln n} + \frac{\ell^2}{2n^2\ln n} + \frac{\ell^2}{(n\ln n)^2} + O\left(\frac{1}{n^3}\right)\right) \\
&\quad\times\left(1 + \frac{\ell}{n\ln n\ln_2 n} + \frac{\ell^2}{2n^2\ln n\ln_2 n} + \frac{\ell^2}{2(n\ln n)^2\ln_2 n} + \frac{\ell^2}{(n\ln n\ln_2 n)^2} + O\left(\frac{1}{n^3}\right)\right) \\
&\quad\times\left(1 + \frac{\ell}{n\ln n\ln_2 n\ln_3 n} + \frac{\ell^2}{2n^2\ln n\ln_2 n\ln_3 n} + \frac{\ell^2}{2(n\ln n)^2\ln_2 n\ln_3 n}\right. \\
&\qquad\left. + \frac{\ell^2}{2(n\ln n\ln_2 n)^2\ln_3 n} + \frac{\ell^2}{(n\ln n\ln_2 n\ln_3 n)^2} + O\left(\frac{1}{n^3}\right)\right) \\
&\quad\times\cdots\times\left(1 + \frac{\ell}{n\ln n\ln_2 n\ln_3 n\cdots\ln_q n} + \frac{\ell^2}{2n^2\ln n\cdots\ln_q n} + \frac{\ell^2}{2(n\ln n)^2\ln_2\cdots n\ln_q n}\right. \\
&\qquad\left. + \cdots + \frac{\ell^2}{2(n\ln n\cdots\ln_{q-1} n)^2\ln_q n} + \frac{\ell^2}{(n\ln n\cdots\ln_q n)^2} + O\left(\frac{1}{n^3}\right)\right).
\end{aligned} \tag{A.14}$$

Finally, gathering the same functional terms and omitting the terms having a higher order of accuracy than is necessary, we obtain the asymptotic decomposition (A.13). □

to Contents, page 1

A.5. Formula for $\alpha^2(n)$

Lemma A.5. *For fixed $q \in \mathbb{N}_0$, the formula*

$$\alpha^2(n) = \frac{4}{3}\sum_{i=0}^{q}\omega_i(n) - \frac{1}{3}\Omega(n) \tag{A.15}$$

holds for all sufficiently large n.

Proof. It is easy to see that

$$\begin{aligned}
\alpha^2(n) = {} & \frac{1}{n^2} + \frac{2}{n^2\ln n} + \frac{2}{n^2\ln n\ln_2 n} + \cdots + \frac{2}{n^2\ln n\ln_2 n\cdots\ln_q n} \\
& + \frac{1}{(n\ln n)^2} + \frac{2}{(n\ln n)^2\ln_2 n} + \cdots + \frac{2}{(n\ln n)^2\ln_2 n\cdots\ln_q n} \\
& + \frac{1}{(n\ln n\ln_2 n)^2} + \frac{2}{(n\ln n\ln_2 n)^2\ln_3 n} + \cdots + \frac{2}{(n\ln n\ln_2 n)^2\cdots\ln_q n} \\
& + \cdots + \frac{1}{(n\ln n\ln_2 n\cdots\ln_q n)^2} \\
= {} & \frac{4}{3}\left(\frac{1}{n^2} + \frac{3}{2n^2\ln n} + \frac{3}{2n^2\ln n\ln_2 n} + \cdots + \frac{3}{2n^2\ln n\ln_2 n\cdots\ln_q n}\right. \\
& + \frac{1}{(n\ln n)^2} + \frac{3}{2(n\ln n)^2\ln_2 n} + \cdots + \frac{3}{2(n\ln n)^2\ln_2 n\cdots\ln_q n} \\
& + \frac{1}{(n\ln n\ln_2 n)^2} + \frac{2}{(n\ln n\ln_2 n)^2\ln_3 n} + \cdots + \frac{2}{(n\ln n\ln_2 n)^2\cdots\ln_q n} \\
& \left. + \cdots + \frac{1}{(n\ln n\ln_2 n\cdots\ln_q n)^2}\right) \\
& - \frac{1}{3}\left(\frac{1}{n^2} + \frac{1}{(n\ \ln n)^2} + \frac{1}{(n\ln n\ln_2 n)^2} + \cdots + \frac{1}{(n\ln n\ln_2 n\cdots\ln_q n)^2}\right) \\
= {} & \frac{4}{3}\sum_{i=0}^{q}\omega_i(n) - \frac{1}{3}\Omega(n).
\end{aligned} \tag{A.16}$$

□

A.6. Asymptotic Decomposition of $V(n+p)$

Lemma A.6. *For fixed $p \in \mathbb{N}$ and $q \in \mathbb{N}_0$, the asymptotic representation*

$$V(n+p) = \varphi(n)\left[k + \Sigma(p)\alpha(n) + S(p)\sum_{i=0}^{q}\omega_i(n)\right] + O\left(\frac{\varphi(n)}{n^3}\right) \tag{A.17}$$

holds for $n \to \infty$.

to Contents, page 1

Proof. It is easy to deduce from formula (A.13) with $\ell = k-p-1, k-p-2, \dots, -p$ that

$$V(n+p) := \varphi(n+p-k+1) + \varphi(n+p-k+2) + \cdots + \varphi(n+p) = \sum_{\ell=-p}^{k-p-1} \varphi(n-\ell) = \varphi(n)$$

$$\begin{aligned}
&\times \sum_{\ell=-p}^{k-p-1} \left(1 + \frac{\ell}{n} + \frac{\ell}{n\ln n} + \frac{\ell}{n\ln n \ln_2 n} + \cdots + \frac{\ell}{n\ln n\ln_2 n\cdots\ln_q n}\right.\\
&\quad + \frac{\ell^2}{n^2} + \frac{3\ell^2}{2n^2\ln n} + \cdots + \frac{3\ell^2}{2n^2\ln n\ln_2 n\cdots\ln_q n} + \frac{\ell^2}{(n\ln n)^2}\\
&\quad + \frac{3\ell^2}{2(n\ln n)^2\ln_2 n} + \frac{3\ell^2}{2(n\ln n)^2\ln_3 n} + \cdots + \frac{3\ell^2}{2(n\ln n)^2\ln_3 n\cdots\ln_q n}\\
&\quad + \frac{\ell^2}{(n\ln n\ln_2 n)^2} + \frac{3\ell^2}{2(n\ln n\ln_2 n)^2\ln_3 n} + \cdots + \frac{3\ell^2}{2(n\ln n\ln_2 n)^2\ln_3 n\cdots\ln_q n}\\
&\quad + \frac{\ell^2}{(n\ln n\ln_2 n\ln_3 n)^2} + \cdots + \frac{3\ell^2}{2(n\ln n\ln_2 n\ln_3 n)^2\ln_4 n\cdots\ln_q n}\\
&\quad + \cdots + \frac{\ell^2}{(n\ln n\ln_2 n\cdots\ln_{q-1} n)^2} + \frac{3\ell^2}{2(n\ln n\ln_2 n\cdots\ln_{q-1} n)^2\ln_q n}\\
&\quad \left.+ \frac{\ell^2}{(n\ln n\ln_2 n\cdots\ln_q n)^2} + O\left(\frac{1}{n^3}\right)\right).
\end{aligned} \tag{A.18}$$

Then

$$\begin{aligned}
V(n+p) &:= \varphi(n)\sum_{\ell=-p}^{k-p-1}\left[1 + \ell\alpha(n) + \ell^2\sum_{i=0}^{q}\omega_i(n) + O\left(\frac{1}{n^3}\right)\right]\\
&= \varphi(n)\left[\sum_{\ell=-p}^{k-p-1} 1 + \alpha(n)\cdot\sum_{\ell=-p}^{k-p-1}\ell + \sum_{\ell=-p}^{k-p-1}\ell^2\cdot\sum_{i=0}^{q}\omega_i(n) + O\left(\frac{1}{n^3}\right)\right]\\
&= \varphi(n)\left[k + \Sigma(p)\alpha(n) + S(p)\cdot\sum_{i=0}^{q}\omega_i(n) + O\left(\frac{1}{n^3}\right)\right]\\
&= \varphi(n)\left[k + \Sigma(p)\alpha(n) + S(p)\sum_{i=0}^{q}\omega_i(n)\right] + O\left(\frac{\varphi(n)}{n^3}\right).
\end{aligned} \tag{A.19}$$

□

to Contents, page 1

A.7. Asymptotic Decomposition of $V^+(n+p)$

Lemma A.7. *For fixed* $p \in \mathbb{N}_0$ *and* $q \in \mathbb{N}_0$, *the asymptotic representation*

$$V^+(n+p) = \varphi(n)\left[k+1+\Sigma^+(p)\alpha(n)+S^+(p)\sum_{i=0}^{q}\omega_i(n)\right]+O\left(\frac{\varphi(n)}{n^3}\right) \tag{A.20}$$

holds for $n \to \infty$.

Proof. By (A.5), (A.13), (A.17), (A.10), and (A.7), we get

$$\begin{aligned}
V^+(n+p) &:= V(n+p)+\varphi(n+p-k)\\
&= \varphi(n)\left[k+\Sigma(p)\alpha(n)+S(p)\sum_{i=0}^{q}\omega_i(n)\right]+O\left(\frac{\varphi(n)}{n^3}\right)+\varphi(n+p-k)\\
&= \varphi(n)\left[k+\Sigma(p)\alpha(n)+S(p)\sum_{i=0}^{q}\omega_i(n)\right]+O\left(\frac{\varphi(n)}{n^3}\right)\\
&\quad+\varphi(n)\Big(1+(k-p)\alpha(n)+(k-p)^2\omega_0(n)+(k-p)^2\omega_1(n)\\
&\qquad+\cdots+(k-p)^2\omega_{q-1}(n)+(k-p)^2\omega_q(n)+O\left(\frac{1}{n^3}\right)\Big)\\
&= \varphi(n)\left[k+1+(\Sigma(p)+(k-p))\alpha(n)+\left(S(p)+(k-p)^2\right)\sum_{i=0}^{q}\omega_i(n)\right]+O\left(\frac{\varphi(n)}{n^3}\right)\\
&= \varphi(n)\left[k+1+\Sigma^+(p)\alpha(n)+S^+(p)\sum_{i=0}^{q}\omega_i(n)\right]+O\left(\frac{\varphi(n)}{n^3}\right).
\end{aligned} \tag{A.21}$$

□

A.8. Formula for $\sum_{p=1}^{k}\Sigma(p)$

Lemma A.8. *For the above sum, the following formula holds:*

$$\sum_{p=1}^{k}\Sigma(p) = -k^2. \tag{A.22}$$

Proof. Using formula (A.9), we get

$$\begin{aligned}
\sum_{p=1}^{k}\Sigma(p) &= \Sigma(1)+\Sigma(2)+\Sigma(3)+\cdots+\Sigma(k)\\
&= \frac{k}{2}\cdot[(k-3)+(k-5)+(k-7)+\cdots+(k-(2k+1))]\\
&= \frac{k}{2}\cdot(-2k) = -k^2.
\end{aligned} \tag{A.23}$$

□

to Contents, page 1

A.9. Formula for $\sum_{p=1}^{k} \Sigma^2(p)$

Lemma A.9. *For the above sum, the following formula holds:*

$$\sum_{p=1}^{k} \Sigma^2(p) = \frac{k^3}{12}\left(k^2 + 11\right). \tag{A.24}$$

Proof. Using formula (A.9), we get

$$\begin{aligned}\sum_{p=1}^{k} \Sigma^2(p) &= \frac{k^2}{4} \sum_{p=1}^{k} (k - 2p - 1)^2 \\ &= \frac{k^2}{4} \cdot \left[(k-3)^2 + (k-5)^2 + (k-7)^2 + \cdots + (k - (2k+1))^2\right].\end{aligned} \tag{A.25}$$

We compute the sum in the square brackets. We use formula (A.12). In our case,

$$\begin{gathered} r = k, \quad u_1 = (k-3)^2, \quad u_2 = (k-5)^2, \quad u_3 = (k-7)^2, \ldots, \quad u_k = (k - 2k - 1)^2 = (k+1)^2, \\ d_1' = u_2 - u_1 = (k-5)^2 - (k-3)^2 = -4k + 16, \\ d_2' = u_3 - u_2 = (k-7)^2 - (k-5)^2 = -4k + 24, \end{gathered} \tag{A.26}$$

the second differences are constant, and

$$d_1'' = d_2' - d_1' = (-4k + 24) - (-4k + 16) = 8. \tag{A.27}$$

Then the sum in the square brackets equals

$$\frac{k!}{(k-1)! \cdot 1!} \cdot (k-3)^2 + \frac{k!(-4)}{(k-2)! \cdot 2!} \cdot (k-4) + \frac{k!}{(k-3)! \cdot 3!} \cdot 8 = \frac{k}{3}\left(k^2 + 11\right), \tag{A.28}$$

and formula (A.24) is proved. □

A.10. Formula for $2\prod_{i,j=0,i>j}^{k} \Sigma(i)\Sigma(j)$

Lemma A.10. *For the above product, the following formula holds:*

$$2\prod_{\substack{i,j=0 \\ i>j}}^{k} \Sigma(i)\Sigma(j) = k^4 - \frac{k^3}{12}\left(k^2 + 11\right). \tag{A.29}$$

to Contents, page 1

Proof. We have

$$2\prod_{\substack{i,j=0\\ i>j}}^{k}\Sigma(i)\ \Sigma(j)=\left(\sum_{p=1}^{k}\Sigma(p)\right)^{2}-\sum_{p=1}^{k}(\Sigma(p))^{2}. \tag{A.30}$$

Then, using formulas (A.22), and (A.24), we get

$$\left(\sum_{p=1}^{k}\Sigma(p)\right)^{2}-\sum_{p=1}^{k}(\Sigma(p))^{2}=\left(-k^{2}\right)^{2}-\frac{k^{3}}{12}\left(k^{2}+11\right)=k^{4}-\frac{k^{3}}{12}\left(k^{2}+11\right). \tag{A.31}$$

□

A.11. Formula for $\sum_{p=0}^{k}\Sigma^{+}(p)$

Lemma A.11. *For the above sum, the following formula holds:*

$$\sum_{p=0}^{k}\Sigma^{+}(p)=0. \tag{A.32}$$

Proof. Using formulas (A.9), (A.10), and (A.22), we get

$$\sum_{p=0}^{k}\Sigma^{+}(p)=\Sigma(0)+\sum_{p=1}^{k}\Sigma(p)+\sum_{p=0}^{k}(k-p)=\frac{k}{2}(k-1)-k^{2}+\frac{k}{2}(k+1)=0. \tag{A.33}$$

□

A.12. Formula for $\sum_{p=0}^{k}(\Sigma^{+}(p))^{2}$

Lemma A.12. *For the above sum, the following formula holds:*

$$\sum_{p=0}^{k}(\Sigma^{+}(p))^{2}=\frac{(k+1)^{2}k}{12}\cdot\left(k^{2}+3k+2\right). \tag{A.34}$$

Proof. Using formula (A.10), we get

$$\sum_{p=0}^{k}(\Sigma^{+}(p))^{2}=\frac{(k+1)^{2}}{4}\left[(k-0)^{2}+(k-2)^{2}+(k-4)^{2}+\cdots+(k-2k)^{2}\right]. \tag{A.35}$$

We compute the sum in the square brackets. We use formula (A.12). In our case,

$$\begin{gathered}r=k+1,\quad u_{1}=k^{2},\quad u_{2}=(k-2)^{2},\quad u_{3}=(k-4)^{2},\ldots,\quad u_{k+1}=(k-2k)^{2}=k^{2},\\ d_{1}'=u_{2}-u_{1}=(k-2)^{2}-k^{2}=-4k+4,\\ d_{2}'=u_{3}-u_{2}=(k-4)^{2}-(k-2)^{2}=-4k+12,\end{gathered} \tag{A.36}$$

to Contents, page 1

the second differences are constant, and

$$d_1'' = d_2' - d_1' = (-4k + 12) - (-4k + 4) = 8. \tag{A.37}$$

Then, the sum in the square brackets equals

$$\frac{(k+1)!}{k! \cdot 1!} \cdot k^2 + \frac{4(k+1)!}{(k-1)! \cdot 2!} \cdot (-k+1) + \frac{(k+1)!}{(k-2)! \cdot 3!} \cdot 8 = \frac{k}{3}\left(k^2 + 3k + 2\right), \tag{A.38}$$

and formula (A.34) is proved. □

A.13. Formula for $2\prod_{i,j=0,\ i>j}^{k} \Sigma^+(i)\Sigma^+(j)$

Lemma A.13. *For the above product, the following formula holds:*

$$2\prod_{\substack{i,j=0\\ i>j}}^{k} \Sigma^+(i)\Sigma^+(j) = -\frac{(k+1)^2 k}{12}\left(k^2 + 3k + 2\right). \tag{A.39}$$

Proof. We have

$$2\prod_{\substack{i,j=0\\ i>j}}^{k} \Sigma^+(i)\Sigma^+(j) = \left(\sum_{p=0}^{k} \Sigma^+(p)\right)^2 - \sum_{p=0}^{k} \left(\Sigma^+(p)\right)^2. \tag{A.40}$$

Then, using formulas (A.32), and (A.34), we get

$$\left(\sum_{p=1}^{k} \Sigma^+(p)\right)^2 - \sum_{p=1}^{k} \left(\Sigma^+(p)\right)^2 = -\sum_{p=1}^{k} \left(\Sigma^+(p)\right)^2 = -\frac{(k+1)^2 k}{12} \cdot \left(k^2 + 3k + 2\right). \tag{A.41}$$

□

A.14. Formula for $S(p)$

Lemma A.14. *For a fixed integer p, the formula*

$$S(p) = \frac{k}{6}\left[2k^2 - 3(1 + 2p)k + \left(6p^2 + 6p + 1\right)\right] \tag{A.42}$$

holds.

to Contents, page 1

Proof. We use formula (A.12). In our case

$$r = k, \quad u_1 = (k-p-1)^2, \dots, \quad u_k = (k-p-k)^2 = p^2,$$
$$d_1' = u_2 - u_1 = (k-p-2)^2 - (k-p-1)^2 = (2k-2p-3)(-1), \tag{A.43}$$
$$d_2' = u_3 - u_2 = (k-p-3)^2 - (k-p-2)^2 = (2k-2p-5)(-1),$$

the second differences are constant, and

$$d_1'' = d_2' - d_1' = (2k-2p-5)(-1) - (2k-2p-3)(-1) = 2. \tag{A.44}$$

Then the formula

$$S(p) = \frac{k!}{(k-1)! \cdot 1!} \cdot (k-p-1)^2 + \frac{k!(-1)}{(k-2)! \cdot 2!} \cdot (2k-2p-3) + \frac{k!}{(k-3)! \cdot 3!} \cdot 2 \tag{A.45}$$

directly follows from (A.12). After some simplification, we get

$$\begin{aligned} S(p) &= k \cdot (k-p-1)^2 - \frac{k(k-1)}{2} \cdot (2k-2p-3) + \frac{k(k-1)(k-2)}{3} \\ &= \frac{k}{6} \cdot \left[6\left(k^2 - 2k(p+1) + (p+1)^2\right) - 3\left(2k^2 - k(2p+5) + (2p+3)\right) + 2\left(k^2 - 3k + 2\right)\right] \\ &= \frac{k}{6}\left[2k^2 - 3(1+2p)k + \left(6p^2 + 6p + 1\right)\right]. \end{aligned} \tag{A.46}$$

Formula (A.42) is proved. □

A.15. Formula for $\sum_{p=1}^{k} S(p)$

Lemma A.15. *For a fixed integer p, the formula*

$$\sum_{p=1}^{k} S(p) = \frac{k}{6}\left(k^3 + 5k\right) \tag{A.47}$$

holds.

Proof. Since, by (A.42),

$$\frac{6}{k} S(p) = 2k^2 - 3(1+2p)k + \left(6p^2 + 6p + 1\right), \tag{A.48}$$

to Contents, page 1

we get

$$\begin{aligned}\frac{6}{k}\sum_{p=1}^{k}S(p) &= 2\sum_{p=1}^{k}k^2 - 3k\sum_{p=1}^{k}(1+2p) + 6\sum_{p=1}^{k}p^2 + 6\sum_{p=1}^{k}p + \sum_{p=1}^{k}1\\ &= 2k^3 - 3k\left(k^2+2k\right) + k\left(2k^2+3k+1\right) + 3\left(k^2+k\right) + k\\ &= k^3+5k.\end{aligned} \tag{A.49}$$

This yields (A.47). □

A.16. Formula for $S^+(p)$

Lemma A.16. *The above expression equals*

$$S^+(p) = \frac{k+1}{6}\left[2k^2 + (-6p+1)k + 6p^2\right]. \tag{A.50}$$

Proof. We have the following:

$$\begin{aligned}S^+(p) &= (k-p)^2 + S(p)\\ &= \left[(k-p)^2 + \frac{k}{6}\left[2k^2 - 3(1+2p)k + \left(6p^2+6p+1\right)\right]\right.\\ &\quad \left. -\frac{k+1}{6}\left[2k^2+(-6p+1)k+6p^2\right]\right] + \frac{k+1}{6}\left[2k^2+(-6p+1)k+6p^2\right]\\ &= \frac{1}{6}\left[6k^2 - 12kp + 6p^2 + k[-4k+6p+1] - 2k^2 + (6p-1)k - 6p^2\right]\\ &\quad + \frac{k+1}{6}\left[2k^2+(-6p+1)k+6p^2\right]\\ &= \frac{k+1}{6}\left[2k^2+(-6p+1)k+6p^2\right].\end{aligned} \tag{A.51}$$

This yields (A.50). □

A.17. Formula for $\sum_{p=0}^{k} S^+(p)$

Lemma A.17. *The above expression equals*

$$\sum_{p=0}^{k}S^+(p) = \frac{(k+1)k}{6}\left(k^2+3k+2\right). \tag{A.52}$$

Proof. Since, by (A.50),

$$\frac{6}{k+1}S^+(p) = 2k^2 + (-6p+1)k + 6p^2, \tag{A.53}$$

to Contents, page 1

we get

$$\begin{aligned}\frac{6}{k+1}\sum_{p=0}^{k}S^{+}(p) &= 2\sum_{p=0}^{k}k^{2}+k\sum_{p=0}^{k}(-6p+1)+6\sum_{p=0}^{k}p^{2}\\ &= 2k^{2}(k+1)+k(-3k(k+1)+(k+1))+k\left(2k^{2}+3k+1\right)\\ &= k^{3}+3k^{2}+2k.\end{aligned} \tag{A.54}$$

This yields (A.52). □

A.18. Formula for $(1/k)\sum_{p=1}^{k}S(p)-(1/(k+1))\sum_{p=0}^{k}S^{+}(p)$

Lemma A.18. *The above expression equals*

$$\frac{1}{k}\sum_{p=1}^{k}S(p)-\frac{1}{k+1}\sum_{p=0}^{k}S^{+}(p)=\frac{1}{2}\cdot\left(-k^{2}+k\right). \tag{A.55}$$

Proof. By (A.47) and (A.50), we obtain

$$\begin{aligned}\frac{1}{k}\sum_{p=1}^{k}S(p)-\frac{1}{k+1}\sum_{p=0}^{k}S^{+}(p) &= \frac{1}{6}\cdot\left(k^{3}+5k\right)-\frac{1}{6}\cdot\left(k^{3}+3k^{2}+2k\right)\\ &= \frac{1}{6}\cdot\left(-3k^{2}+3k\right)=\frac{1}{2}\cdot\left(-k^{2}+k\right).\end{aligned} \tag{A.56}$$

This yields (A.55). □

A.19. Asymptotic Decomposition of $\prod_{p=1}^{k}V(n+p)$

Lemma A.19. *For a fixed $q\in\mathbb{N}_0$, the asymptotic representation*

$$\begin{aligned}\prod_{p=1}^{k}V(n+p) &= k^{k}\varphi^{k}(n)\left[1-k\alpha(n)-\frac{k}{24}\left(k^{2}-12k+11\right)\alpha^{2}(n)+\frac{k}{6}\left(k^{2}+5\right)\sum_{i=0}^{q}\omega_{i}(n)\right]\\ &\quad+O\left(\frac{\varphi^{k}(n)}{n^{3}}\right)\end{aligned} \tag{A.57}$$

holds for $n\rightarrow\infty$.

to Contents, page 1

Proof. Using formula (A.17), we get

$$\prod_{p=1}^{k} V(n+p) = \prod_{p=1}^{k}\left[\varphi(n)\left[k+\Sigma(p)\alpha(n)+S(p)\sum_{i=0}^{q}\omega_i(n)\right]+O\left(\frac{\varphi(n)}{n^3}\right)\right]$$

$$= \varphi^k(n)\left[k^k + k^{k-1}\alpha(n)\sum_{i=1}^{k}\Sigma(i) + k^{k-2}\alpha^2(n)\prod_{\substack{i,j=0\\ i>j}}^{k}\Sigma(i)\ \Sigma(j) + k^{k-1}\sum_{i=1}^{k}S(i)\sum_{j=0}^{q}\omega_j(n)\right]$$

$$+ O\left(\frac{\varphi^k(n)}{n^3}\right) = (*). \tag{A.58}$$

Now, by (A.22), (A.29), and (A.47)

$$(*) = \varphi^k(n)\left[k^k + k^{k-1}(-k)^2\alpha(n) + \frac{1}{2}k^{k-2}\left(k^4 - \frac{k^3}{12}\left(k^2+11\right)\right)\alpha^2(n)\right.$$

$$\left.+\frac{1}{6}k^{k-1}k\left(k^3+5k\right)\sum_{j=0}^{q}\omega_j(n)\right] + O\left(\frac{\varphi^k(n)}{n^3}\right) \tag{A.59}$$

$$= k^k\varphi^k(n)\left[1 - k\alpha(n) - \frac{k}{24}\left(k^2-12k+11\right)\alpha^2(n)\right.$$

$$\left.+\frac{k}{6}\left(k^2+5\right)\sum_{j=0}^{q}\omega_j(n)\right] + O\left(\frac{\varphi^k(n)}{n^3}\right).$$

□

A.20. Asymptotic Decomposition of $\prod_{p=0}^{k} V^+(n+p)$

Lemma A.20. *For a fixed* $q \in \mathbb{N}_0$*, the asymptotic representation*

$$\prod_{p=0}^{k} V^+(n+p) = (k+1)^{k+1}\varphi^{k+1}(n)\left[1 - \frac{k}{24}\left(k^2+3k+2\right)\alpha^2(n) + \frac{k}{6}\left(k^2+3k+2\right)\sum_{i=0}^{q}\omega_i(n)\right]$$

$$+ O\left(\frac{\varphi^{k+1}(n)}{n^3}\right) \tag{A.60}$$

holds for $n \to \infty$.

to Contents, page 1

Proof. Using formula (A.20), we get

$$\begin{aligned}
\prod_{p=0}^{k} V^{+}(n+p) &= \prod_{p=0}^{k}\left[\varphi(n)\left[k+1+\Sigma^{+}(p)\alpha(n)+S^{+}(p)\sum_{i=0}^{q}\omega_i(n)\right]+O\left(\frac{\varphi(n)}{n^3}\right)\right] \\
&= \varphi^{k+1}(n)\Bigg[(k+1)^{k+1}+(k+1)^{k}\alpha(n)\sum_{i=0}^{k}\Sigma^{+}(i) \\
&\quad +(k+1)^{k-1}\alpha^2(n)\prod_{\substack{i,j=0\\ i>j}}^{k}\Sigma^{+}(i)\Sigma^{+}(j) \\
&\quad +(k+1)^{k}\sum_{i=0}^{k}S^{+}(i)\sum_{j=0}^{q}\omega_j(n)\Bigg]+O\left(\frac{\varphi^{k+1}(n)}{n^3}\right)=(*).
\end{aligned} \tag{A.61}$$

Now, by (A.32), (A.39), and (A.52), we derive

$$\begin{aligned}
(*) &= \varphi^{k+1}(n)\Bigg[(k+1)^{k+1}-(k+1)^{k-1}\frac{(k+1)^2k}{24}\left(k^2+3k+2\right)\alpha^2(n) \\
&\quad +(k+1)^{k}\frac{(k+1)k}{6}\left(k^2+3k+2\right)\sum_{j=0}^{q}\omega_j(n)\Bigg]+O\left(\frac{\varphi^{k+1}(n)}{n^3}\right) \\
&= (k+1)^{k+1}\varphi^{k+1}(n)\left[1-\frac{k}{24}\left(k^2+3k+2\right)\alpha^2(n)+\frac{k}{6}\left(k^2+3k+2\right)\sum_{j=0}^{q}\omega_j(n)\right] \\
&\quad +O\left(\frac{\varphi^{k+1}(n)}{n^3}\right).
\end{aligned} \tag{A.62}$$

□

Acknowledgment

This research was supported by Grants P201/10/1032 and P201/11/0768 of the Czech Grant Agency (Prague), by the Council of Czech Government MSM 00216 30503 and by the project FEKT-S-11-2(921) Brno University of Technology.

References

[1] R. P. Agarwal, M. Bohner, and W.-T. Li, *Nonoscillation and Oscillation: Theory for Functional Differential Equations*, vol. 267 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 2004.

to Contents, page 1

[2] R. P. Agarwal and A. Zafer, "Oscillation criteria for second-order forced dynamic equations with mixed nonlinearities," *Advances in Difference Equations*, vol. 2009, Article ID 938706, 2009 pages, 2009.

[3] J. Baštinec, L. Berezansky, J. Diblík, and Z. Šmarda, "On the critical case in oscillation for differential equations with a single delay and with several delays," *Abstract and Applied Analysis*, vol. 2010, Article ID 417869, 20 pages, 2010.

[4] J. Baštinec, J. Diblík, and Z. Šmarda, "Existence of positive solutions of discrete linear equations with a single delay," *Journal of Difference Equations and Applications*, vol. 16, no. 9, pp. 1047–1056, 2010.

[5] J. Baštinec, J. Diblík, and Z. Šmarda, "Oscillation of solutions of a linear second-order discretedelayed equation," *Advances in Difference Equations*, vol. 2010, Article ID 693867, 12 pages, 2010.

[6] L. Berezansky and E. Braverman, "On existence of positive solutions for linear difference equations with several delays," *Advances in Dynamical Systems and Applications*, vol. 1, no. 1, pp. 29–47, 2006.

[7] L. Berezansky and E. Braverman, "Oscillation of a logistic difference equation with several delays," *Advances in Difference Equations*, vol. 2006, Article ID 82143, 12 pages, 2006.

[8] M. Bohner, B. Karpuz, and Ö. Öcalan, "Iterated oscillation criteria for delay dynamic equations of first order," *Advances in Difference Equations*, vol. 2008, Article ID 458687, 12 pages, 2008.

[9] G. E. Chatzarakis, R. Koplatadze, and I. P. Stavroulakis, "Oscillation criteria of first order linear difference equations with delay argument," *Nonlinear Analysis*, vol. 68, no. 4, pp. 994–1005, 2008.

[10] J. Diblík, "Positive and oscillating solutions of differential equations with delay in critical case," *Journal of Computational and Applied Mathematics*, vol. 88, no. 1, pp. 185–202, 1998.

[11] J. Diblík and N. Koksch, "Positive solutions of the equation $\dot{x}(t) = -c(t)x(t-\tau)$ in the critical case," *Journal of Mathematical Analysis and Applications*, vol. 250, no. 2, pp. 635–659, 2000.

[12] J. Diblík, Z. Svoboda, and Z. Šmarda, "Explicit criteria for the existence of positive solutions for a scalar differential equation with variable delay in the critical case," *Computers & Mathematics with Applications*, vol. 56, no. 2, pp. 556–564, 2008.

[13] Y. Domshlak, "Oscillation properties of discrete difference inequalities and equations: the new approach," in *Functional-Differential Equations*, vol. 1 of *Functional Differential Equations Israel Seminar*, pp. 60–82, Coll. Judea Samaria, Ariel, Israel, 1993.

[14] Y. Domshlak and I. P. Stavroulakis, "Oscillations of first-order delay differential equations in a critical state," *Applicable Analysis*, vol. 61, no. 3-4, pp. 359–371, 1996.

[15] B. Dorociaková and R. Olach, "Existence of positive solutions of delay differential equations," *Tatra Mountains Mathematical Publications*, vol. 43, pp. 63–70, 2009.

[16] S. Elaydi, *An Introduction to Difference Equations*, Undergraduate Texts in Mathematics, Springer, New York, NY, USA, 3rd edition, 2005.

[17] Á. Elbert and I. P. Stavroulakis, "Oscillation and nonoscillation criteria for delay differential equations," *Proceedings of the American Mathematical Society*, vol. 123, no. 5, pp. 1503–1510, 1995.

[18] L. H. Erbe and B. G. Zhang, "Oscillation of discrete analogues of delay equations," *Differential and Integral Equations*, vol. 2, no. 3, pp. 300–309, 1989.

[19] I. Győri and G. Ladas, *Oscillation Theory of Delay Differential Equations*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, NY, USA, 1991.

[20] L. Hanuštiaková and R. Olach, "Nonoscillatory bounded solutions of neutral differential systems," *Nonlinear Analysis*, vol. 68, no. 7, pp. 1816–1824, 2008.

[21] L. K. Kikina and I. P. Stavroulakis, "A survey on the oscillation of solutions of first order delay difference equations," *Cubo*, vol. 7, no. 2, pp. 223–236, 2005.

[22] G. Ladas, Ch. G. Philos, and Y. G. Sficas, "Sharp conditions for the oscillation of delay difference equations," *Journal of Applied Mathematics and Simulation*, vol. 2, no. 2, pp. 101–111, 1989.

[23] R. Medina and M. Pituk, "Nonoscillatory solutions of a second-order difference equation of Poincaré type," *Applied Mathematics Letters*, vol. 22, no. 5, pp. 679–683, 2009.

[24] V. E. Slyusarchuk, "Necessary and sufficient conditions for the oscillation of solutions of nonlinear differential equations with impulse action in a Banach space," *Ukrainian Mathematical Journal*, vol. 51, no. 1, pp. 98–109, 1999.

[25] I. P. Stavroulakis, "Oscillation criteria for first order delay difference equations," *Mediterranean Journal of Mathematics*, vol. 1, no. 2, pp. 231–240, 2004.

[26] X. H. Tang and J. S. Yu, "Oscillations of delay difference equations in a critical state," *Applied Mathematics Letters*, vol. 13, no. 2, pp. 9–15, 2000.

[27] X. H. Tang and J. S. Yu, "Oscillation of delay difference equations," *Computers & Mathematics with Applications*, vol. 37, pp. 11–20, 1999.

[28] B. G. Zhang and S. S. Cheng, "Oscillation criteria and comparison theorems for delay difference equations," *Fasciculi Mathematici*, no. 25, pp. 13–32, 1995.

to Contents, page 1

[29] Y. Zhou and B. G. Zhang, "Oscillations of delay difference equations in a critical state," *Computers & Mathematics with Applications*, vol. 39, no. 7-8, pp. 71–80, 2000.

[30] R. Y. Zhang, Z. C. Wang, and X. Zou, "Oscillation of delay difference equations: near a critical case," *Dynamics of Continuous, Discrete & Impulsive Systems. Series A. Mathematical Analysis*, vol. 9, no. 1, pp. 89–103, 2002.

[31] L. Berg, "On the asymptotics of nonlinear difference equations," *Zeitschrift für Analysis und Ihre Anwendungen*, vol. 21, no. 4, pp. 1061–1074, 2002.

[32] L. Berg, "Inclusion theorems for non-linear difference equations with applications," *Journal of Difference Equations and Applications*, vol. 10, no. 4, pp. 399–408, 2004.

[33] L. Berg, "On the asymptotics of the difference equation $x_{n+1} = x_n x_{n-2} - 1$," *Journal of Difference Equations and Applications*, vol. 14, no. 1, pp. 105–108, 2008.

[34] B. Iričanin and S. Stević, "Eventually constant solutions of a rational difference equation," *Applied Mathematics and Computation*, vol. 215, no. 2, pp. 854–856, 2009.

[35] C. M. Kent, W. Kosmala, and S. Stevič, "On the difference equation," *Abstract and Applied Analysis*, vol. 2011, Article ID 815285, 15 pages, 2011.

[36] S. Stević, "On the recursive sequence $x_{n+1} = x_{n-1}/g(x_n)$," *Taiwanese Journal of Mathematics*, vol. 6, no. 3, pp. 405–414, 2002.

[37] S. Stević, "Global stability and asymptotics of some classes of rational difference equations," *Journal of Mathematical Analysis and Applications*, vol. 316, no. 1, pp. 60–68, 2006.

[38] S. Stević, "On monotone solutions of some classes of difference equations," *Discrete Dynamics in Nature and Society*, vol. 2006, Article ID 53890, 9 pages, 2006.

[39] S. Stević, "On positive solutions of a $(k+1)$-th order difference equation," *Applied Mathematics Letters*, vol. 19, no. 5, pp. 427–431, 2006.

[40] S. Stević, "Asymptotics of some classes of higher-order difference equations," *Discrete Dynamics in Nature and Society*, vol. 2007, Article ID 56813, 20 pages, 2007.

[41] S. Stević, "Existence of nontrivial solutions of a rational difference equation," *Applied Mathematics Letters*, vol. 20, no. 1, pp. 28–31, 2007.

[42] S. Stevich, "Nontrivial solutions of higher-order rational difference equations," *Rossiĭskaya Akademiya Nauk*, vol. 84, no. 5, pp. 772–780, 2008.

[43] H. B. Dwight, *Tables of Integrals and Other Mathematical Data*, The Macmillan Company, New York, NY, USA, 4th edition, 1961.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 714248, 18 pages
doi:10.1155/2011/714248

Research Article

A Two-Species Cooperative Lotka-Volterra System of Degenerate Parabolic Equations

Jiebao Sun, Dazhi Zhang, and Boying Wu

Department of Mathematics, Harbin Institute of Technology, Harbin 150001, China

Correspondence should be addressed to Jiebao Sun, sunjiebao@126.com

Received 13 December 2010; Revised 13 February 2011; Accepted 24 February 2011

Academic Editor: Elena Braverman

We consider a cooperating two-species Lotka-Volterra model of degenerate parabolic equations. We are interested in the coexistence of the species in a bounded domain. We establish the existence of global generalized solutions of the initial boundary value problem by means of parabolic regularization and also consider the existence of the nontrivial time-periodic solution for this system.

1. Introduction

In this paper, we consider the following two-species cooperative system:

$$u_t = \Delta u^{m_1} + u^{\alpha}(a - bu + cv), \quad (x,t) \in \Omega \times \mathbb{R}_+, \tag{1.1}$$

$$v_t = \Delta v^{m_2} + v^{\beta}(d + eu - fv), \quad (x,t) \in \Omega \times \mathbb{R}_+, \tag{1.2}$$

$$u(x,t) = 0, \quad v(x,t) = 0, \quad (x,t) \in \partial\Omega \times \mathbb{R}_+, \tag{1.3}$$

$$u(x,0) = u_0(x), \quad v(x,0) = v_0(x), \quad x \in \Omega, \tag{1.4}$$

where $m_1, m_2 > 1$, $0 < \alpha < m_1$, $0 < \beta < m_2$, $1 \le (m_1 - \alpha)(m_2 - \beta)$, $a = a(x,t)$, $b = b(x,t)$, $c = c(x,t)$, $d = d(x,t)$, $e = e(x,t)$, $f = f(x,t)$ are strictly positive smooth functions and periodic in time with period $T > 0$ and $u_0(x)$ and $v_0(x)$ are nonnegative functions and satisfy $u_0^{m_1}, v_0^{m_2} \in W_0^{1,2}(\Omega)$.

In dynamics of biological groups, the system (1.1)-(1.2) can be used to describe the interaction of two biological groups. The diffusion terms Δu^{m_1} and Δv^{m_2} represent the effect of dispersion in the habitat, which models a tendency to avoid crowding and the speed of the diffusion is rather slow. The boundary conditions (1.3) indicate that the habitat is surrounded by a totally hostile environment. The functions u and v represent the spatial densities of the species at time t and a, d are their respective net birth rate. The functions b and f are intra-specific competitions, whereas c and e are those of interspecific competitions.

As famous models for dynamics of population, two-species cooperative systems like (1.1)-(1.2) have been studied extensively, and there have been many excellent results, for detail one can see [1–6] and references therein. As a special case, men studied the following two-species Lotka-Volterra cooperative system of ODEs:

$$\begin{aligned} u'(t) &= u(t)(a(t) - b(t)u(t) + c(t)v(t)),\\ v'(t) &= v(t)(d(t) + e(t)u(t) - f(t)v(t)). \end{aligned} \tag{1.5}$$

For this system, Lu and Takeuchi [7] studied the stability of positive periodic solution and Cui [1] discussed the persistence and global stability of it.

When $m_1 = m_2 = \alpha = \beta = 1$, from (1.1)-(1.2) we get the following classical cooperative system:

$$\begin{aligned} u_t &= \Delta u + u(a - bu + cv),\\ v_t &= \Delta v + v(d + eu - fv). \end{aligned} \tag{1.6}$$

For this system, Lin et al. [5] showed the existence and asymptotic behavior of T-periodic solutions when a, b, c, e, d, f are all smooth positive and periodic in time with period $T > 0$. When a, b, c, e, d, f are all positive constants, Pao [6] proved that the Dirichlet boundary value problem of this system admits a unique solution which is uniformly bounded when $ce < bf$, while the blowup solutions are possible when the two species are strongly mutualistic ($ce > bf$). For the homogeneous Neumann boundary value problem of this system, Lou et al. [4] proved that the solution will blow up in finite time under a sufficient condition on the initial data. When $c = e = 0$ and $\alpha = \beta = 1$, from (1.1) we get the single degenerate equation

$$u_t = \Delta u^m + u(a - bu). \tag{1.7}$$

For this equation, Sun et al. [8] established the existence of nontrivial nonnegative periodic solutions by monotonicity method and showed the attraction of nontrivial nonnegative periodic solutions.

In the recent years, much attention has been paid to the study of periodic boundary value problems for parabolic systems; for detail one can see [9–15] and the references therein. Furthermore, many researchers studied the periodic boundary value problem for degenerate parabolic systems, such as [16–19]. Taking into account the impact of periodic factors on the species dynamics, we are also interested in the existence of the nontrivial periodic solutions of the cooperative system (1.1)-(1.2). In this paper, we first show the existence of the global generalized solution of the initial boundary value problem (1.1)–(1.4). Then under the condition that

$$b_l f_l > c_M e_M, \tag{1.8}$$

to Contents, page 1

where $f_M = \sup\{f(x,t) \mid (x,t) \in \Omega \times \mathbb{R}\}$, $f_l = \inf\{f(x,t) \mid (x,t) \in \Omega \times \mathbb{R}\}$, we show that the generalized solution is uniformly bounded. At last, by the method of monotone iteration, we establish the existence of the nontrivial periodic solutions of the system (1.1)-(1.2), which follows from the existence of a pair of large periodic supersolution and small periodic subsolution. At last, we show the existence and the attractivity of the maximal periodic solution.

Our main efforts center on the discussion of generalized solutions, since the regularity follows from a quite standard approach. Hence we give the following definition of generalized solutions of the problem (1.1)–(1.4).

Definition 1.1. A nonnegative and continuous vector-valued function (u,v) is said to be a generalized solution of the problem (1.1)–(1.4) if, for any $0 \le \tau < T$ and any functions $\varphi_i \in C^1(\overline{Q}_\tau)$ with $\varphi_i|_{\partial\Omega\times[0,\tau)} = 0$ $(i=1,2)$, $\nabla u^{m_1}, \nabla v^{m_2} \in L^2(Q_\tau)$, $\partial u^{m_1}/\partial t$, $\partial v^{m_2}/\partial t \in L^2(Q_\tau)$ and

$$\iint_{Q_\tau} u\frac{\partial\varphi_1}{\partial t} - \nabla u^{m_1}\nabla\varphi_1 + u^\alpha(a - bu + cv)\varphi_1 dx\,dt = \int_\Omega u(x,\tau)\varphi_1(x,\tau)dx - \int_\Omega u_0(x)\varphi_1(x,0)dx,$$
$$\iint_{Q_\tau} v\frac{\partial\varphi_2}{\partial t} - \nabla v^{m_2}\nabla\varphi_2 + v^\beta(d + eu - fv)\varphi_2 dx\,dt = \int_\Omega v(x,\tau)\varphi_2(x,\tau)dx - \int_\Omega v_0(x)\varphi_2(x,0)dx, \tag{1.9}$$

where $Q_\tau = \Omega \times (0,\tau)$.

Similarly, we can define a weak supersolution $(\overline{u},\overline{v})$ (subsolution $(\underline{u},\underline{v})$) if they satisfy the inequalities obtained by replacing "=" with "≤" ("≥") in (1.3), (1.4), and (1.9) and with an additional assumption $\varphi_i \ge 0$ $(i = 1,\ 2)$.

Definition 1.2. A vector-valued function (u,v) is said to be a T-periodic solution of the problem (1.1)–(1.3) if it is a solution in $[0,T]$ such that $u(\cdot,0) = u(\cdot,T)$, $v(\cdot,0) = v(\cdot,T)$ in Ω. A vector-valued function $(\overline{u},\overline{v})$ is said to be a T-periodic supersolution of the problem (1.1)–(1.3) if it is a supersolution in $[0,T]$ such that $\overline{u}(\cdot,0) \ge \overline{u}(\cdot,T)$, $\overline{v}(\cdot,0) \ge \overline{v}(\cdot,T)$ in Ω. A vector-valued function $(\underline{u},\underline{v})$ is said to be a T-periodic subsolution of the problem (1.1)–(1.3), if it is a subsolution in $[0,T]$ such that $\underline{u}(\cdot,0) \le \underline{u}(\cdot,T)$, $\underline{v}(\cdot,0) \le \underline{v}(\cdot,T)$ in Ω.

This paper is organized as follows. In Section 2, we show the existence of generalized solutions to the initial boundary value problem and also establish the comparison principle. Section 3 is devoted to the proof of the existence of the nonnegative nontrivial periodic solutions by using the monotone iteration technique.

2. The Initial Boundary Value Problem

To solve the problem (1.1)–(1.4), we consider the following regularized problem:

$$\frac{\partial u_\varepsilon}{\partial t} = \operatorname{div}\left(\left(m u_\varepsilon^{m_1-1} + \varepsilon\right)\nabla u_\varepsilon\right) + u_\varepsilon^\alpha(a - bu_\varepsilon + cv_\varepsilon), \quad (x,t) \in Q_T, \tag{2.1}$$

$$\frac{\partial v_\varepsilon}{\partial t} = \operatorname{div}\left(\left(m v_\varepsilon^{m_2-1} + \varepsilon\right)\nabla v_\varepsilon\right) + v_\varepsilon^\beta(d + eu_\varepsilon - fv_\varepsilon), \quad (x,t) \in Q_T, \tag{2.2}$$

$$u_\varepsilon(x,t) = 0, \quad v_\varepsilon(x,t) = 0, \quad (x,t) \in \partial\Omega \times (0,T), \tag{2.3}$$

$$u_\varepsilon(x,0) = u_{0\varepsilon}(x), \quad v_\varepsilon(x,0) = v_{0\varepsilon}(x), \quad x \in \Omega, \tag{2.4}$$

to Contents, page 1

where $Q_T = \Omega\times(0,T)$, $0<\varepsilon<1$, $u_{0\varepsilon}, v_{0\varepsilon} \in C_0^\infty(\Omega)$ are nonnegative bounded smooth functions and satisfy

$$0 \le u_{0\varepsilon} \le \|u_0\|_{L^\infty(\Omega)}, \qquad 0 \le v_{0\varepsilon} \le \|v_0\|_{L^\infty(\Omega)},$$
$$u_{0\varepsilon}^{m_1} \longrightarrow u_0^{m_1}, \quad v_{0\varepsilon}^{m_2} \longrightarrow v_0^{m_2}, \quad \text{in } W_0^{1,2}(\Omega) \text{ as } \varepsilon \longrightarrow 0. \tag{2.5}$$

The standard parabolic theory (cf. [20, 21]) shows that (2.1)–(2.4) admits a nonnegative classical solution $(u_\varepsilon, v_\varepsilon)$. So, the desired solution of the problem (1.1)–(1.4) will be obtained as a limit point of the solutions $(u_\varepsilon, v_\varepsilon)$ of the problem (2.1)–(2.4). In the following, we show some important uniform estimates for $(u_\varepsilon, v_\varepsilon)$.

Lemma 2.1. *Let $(u_\varepsilon, v_\varepsilon)$ be a solution of the problem* (2.1)–(2.4).

(1) *If $1 < (m_1-\alpha)(m_2-\beta)$, then there exist positive constants r and s large enough such that*

$$\frac{1}{m_2-\beta} < \frac{m_1+r-1}{m_2+s-1} < m_1-\alpha, \tag{2.6}$$

$$\|u_\varepsilon\|_{L^r(Q_T)} \le C, \qquad \|v_\varepsilon\|_{L^s(Q_T)} \le C, \tag{2.7}$$

where C is a positive constant only depending on $m_1, m_2, \alpha, \beta, r, s$, $|\Omega|$, and T.

(2) *If $1 = (m_1-\alpha)(m_2-\beta)$, then* (2.7) *also holds when $|\Omega|$ is small enough.*

Proof. Multiplying (2.1) by u_ε^{r-1} $(r>1)$ and integrating over Ω, we have that

$$\int_\Omega \frac{\partial u_\varepsilon^r}{\partial t} dx = -\frac{4r(r-1)m_1}{(m_1+r-1)^2}\int_\Omega \left|\nabla u_\varepsilon^{(m_1+r-1)/2}\right|^2 dx + r\int_\Omega u_\varepsilon^{\alpha+r-1}(a-bu_\varepsilon+cv_\varepsilon)dx. \tag{2.8}$$

By Poincaré's inequality, we have that

$$K\int_\Omega u_\varepsilon^{m_1+r-1}dx \le \int_\Omega \left|\nabla u_\varepsilon^{(m_1+r-1)/2}\right|^2 dx, \tag{2.9}$$

where K is a constant depending only on $|\Omega|$ and N and becomes very large when the measure of the domain Ω becomes small. Since $\alpha < m_1$, Young's inequality shows that

$$au_\varepsilon^{\alpha+r-1} \le \frac{Kr(r-1)m_1}{(m_1+r-1)^2}u_\varepsilon^{m_1+r-1} + CK^{-(\alpha+r-1)/(m_1-\alpha)},$$
$$cu_\varepsilon^{\alpha+r-1}v_\varepsilon \le \frac{Kr(r-1)m_1}{(m_1+r-1)^2}u_\varepsilon^{m_1+r-1} + CK^{-(\alpha+r-1)/(m_1-\alpha)}v_\varepsilon^{(m_1+r-1)/(m_1-\alpha)}. \tag{2.10}$$

to Contents, page 1

For convenience, here and below, C denotes a positive constant which is independent of ε and may take different values on different occasions. Complying (2.8) with (2.9) and (2.10), we obtain

$$\int_{\Omega} \frac{\partial u_{\varepsilon}^{r}}{\partial t} dx \leq -\frac{2Kr(r-1)m_1}{(m_1+r-1)^2} \int_{\Omega} u_{\varepsilon}^{m_1+r-1} dx + CK^{-(\alpha+r-1)/(m_1-\alpha)} \int_{\Omega} v_{\varepsilon}^{(m_1+r-1)/(m_1-\alpha)} dx + CK^{-(\alpha+r-1)/(m_1-\alpha)}. \tag{2.11}$$

As a similar argument as above, for v_{ε} and positive constant $s > 1$, we have that

$$\int_{\Omega} \frac{\partial v_{\varepsilon}^{s}}{\partial t} dx \leq -\frac{2Ks(s-1)m_2}{(m_2+s-1)^2} \int_{\Omega} v_{\varepsilon}^{m_2+s-1} dx + CK^{-(\beta+s-1)/(m_2-\beta)} \int_{\Omega} u_{\varepsilon}^{(m_2+s-1)/(m_2-\beta)} dx + CK^{-(\beta+s-1)/(m_2-\beta)}. \tag{2.12}$$

Thus we have that

$$\begin{aligned} \int_{\Omega} \left(\frac{\partial u_{\varepsilon}^{r}}{\partial t} + \frac{\partial v_{\varepsilon}^{s}}{\partial t} \right) dx \leq & -\frac{2Kr(r-1)m_1}{(m_1+r-1)^2} \int_{\Omega} u_{\varepsilon}^{m_1+r-1} dx + CK^{-(\beta+s-1)/(m_2-\beta)} \int_{\Omega} u_{\varepsilon}^{(m_2+s-1)/(m_2-\beta)} dx \\ & - \frac{2Ks(s-1)m_2}{(m_2+s-1)^2} \int_{\Omega} v_{\varepsilon}^{m_2+s-1} dx + CK^{-(\alpha+r-1)/(m_1-\alpha)} \int_{\Omega} v_{\varepsilon}^{(m_1+r-1)/(m_1-\alpha)} dx \\ & + CK^{-(\alpha+r-1)/(m_1-\alpha)} + CK^{-(\beta+s-1)/(m_2-\beta)}. \end{aligned} \tag{2.13}$$

For the case of $1 < (m_1-\alpha)(m_2-\beta)$, there exist r, s large enough such that

$$\frac{1}{m_1-\alpha} < \frac{m_2+s-1}{m_1+r-1} < m_2-\beta. \tag{2.14}$$

By Young's inequality, we have that

$$\begin{aligned} \int_{\Omega} u_{\varepsilon}^{(m_2+s-1)/(m_2-\beta)} dx &\leq \frac{r(r-1)m_1 K^{(m_2+s-1)/(m_2-\beta)}}{C(m_1+r-1)^2} \int_{\Omega} u_{\varepsilon}^{m_1+r-1} dx + CK^{-\gamma_1}, \\ \int_{\Omega} v_{\varepsilon}^{(m_1+r-1)/(m_1-\alpha)} dx &\leq \frac{s(s-1)m_2 K^{(m_1+r-1)/(m_1-\alpha)}}{C(m_2+s-1)^{p_2}} \int_{\Omega} v_{\varepsilon}^{m_2+s-1} dx + CK^{-\gamma_2}, \end{aligned} \tag{2.15}$$

where

$$\begin{aligned} \gamma_1 &= \frac{(m_2+s-1)^2}{[m_2-\beta]\left[(m_2-\beta)(m_1+r-1)-(m_2+s-1)\right]}, \\ \gamma_2 &= \frac{(m_1+r-1)^2}{[m_1-\alpha]\left[(m_1-\alpha)(m_2+s-1)-(m_1+r-1)\right]}. \end{aligned} \tag{2.16}$$

to Contents, page 1

Together with (2.13), we have that

$$\int_\Omega \left(\frac{\partial u_\varepsilon^r}{\partial t} + \frac{\partial v_\varepsilon^s}{\partial t}\right)dx \le -K\int_\Omega \left(u_\varepsilon^{m_1+r-1} + v_\varepsilon^{m_2+s-1}\right)dx + C\left(K^{-\theta_1} + K^{-\theta_2}\right) + CK^{-(\alpha+r-1)/(m_1-\alpha)} + CK^{-(\beta+s-1)/(m_2-\beta)}, \tag{2.17}$$

where

$$\theta_1 = \frac{(m_2+s-1)+(m_1+r-1)(\beta+s-1)}{(m_2-\beta)(m_1+r-1)-(m_2+s-1)}, \qquad \theta_2 = \frac{(m_1+r-1)+(m_2+s-1)(\alpha+r-1)}{(m_1-\alpha)(m_2+s-1)-(m_1+r-1)}. \tag{2.18}$$

Furthermore, by Hölder's and Young's inequalities, from (2.17) we obtain

$$\int_\Omega \left(\frac{\partial u_\varepsilon^r}{\partial t} + \frac{\partial v_\varepsilon^s}{\partial t}\right)dx \le -K\int_\Omega (u_\varepsilon^r + v_\varepsilon^s)dx + C\left(K^{-\theta_1} + K^{-\theta_2}\right) + 2K|\Omega| + CK^{-(\alpha+r-1)/(m_1-\alpha)} + CK^{-(\beta+s-1)/(m_2-\beta)}. \tag{2.19}$$

Then by Gronwall's inequality, we obtain

$$\int_\Omega (u_\varepsilon^r + v_\varepsilon^s)dx \le C. \tag{2.20}$$

Now we consider the case of $1 = (m_1-\alpha)(m_2-\beta)$. It is easy to see that there exist positive constants r, s large enough such that

$$\frac{1}{m_1-\alpha} = \frac{m_2+s-1}{m_1+r-1} = m_2-\beta. \tag{2.21}$$

Due to the continuous dependence of K upon $|\Omega|$ in (2.9), from (2.13) we have that

$$\int_\Omega \left(\frac{\partial u_\varepsilon^r}{\partial t} + \frac{\partial v_\varepsilon^s}{\partial t}\right)dx \le -K\int_\Omega \left(u_\varepsilon^{m_1+r-1} + v_\varepsilon^{m_2(p_2-1)+s-1}\right)dx + C \tag{2.22}$$

when $|\Omega|$ is small enough. Then by Young's and Gronwall's inequalities we can also obtain (2.20), and thus we complete the proof of this lemma. □

Taking $u_\varepsilon^{m_1}, v_\varepsilon^{m_2}$ as the test functions, we can easily obtain the following lemma.

Lemma 2.2. *Let* $(u_\varepsilon, v_\varepsilon)$ *be a solution of* (2.1)–(2.4); *then*

$$\iint_{Q_T} |\nabla u_\varepsilon^{m_1}|^2 dx\,dt \le C, \qquad \iint_{Q_T} |\nabla v_\varepsilon^{m_2}|^2 dx\,dt \le C, \tag{2.23}$$

where C is a positive constant independent of ε.

to Contents, page 1

Lemma 2.3. *Let* $(u_\varepsilon, v_\varepsilon)$ *be a solution of* (2.1)–(2.4), *then*

$$\|u_\varepsilon\|_{L^\infty(Q_T)} \le C, \qquad \|v_\varepsilon\|_{L^\infty(Q_T)} \le C, \tag{2.24}$$

where C *is a positive constant independent of* ε.

Proof. For a positive constant $k > \|u_{0\varepsilon}\|_{L^\infty(\Omega)}$, multiplying (2.1) by $(u_\varepsilon - k)_+^{m_1}\chi_{[t_1,t_2]}$ and integrating the results over Q_T, we have that

$$\begin{aligned}
&\frac{1}{m_1+1}\iint_{Q_T}\frac{\partial(u_\varepsilon-k)_+^{m_1+1}\chi_{[t_1,t_2]}}{\partial t}dx\,dt+\iint_{Q_T}\left|\nabla(u_\varepsilon-k)_+^{m_1}\chi_{[t_1,t_2]}\right|^2dx\,dt\\
&\quad\le\iint_{Q_T}au_\varepsilon^{\alpha+m_1}(a+cv_\varepsilon)dx\,dt,
\end{aligned}\tag{2.25}$$

where $s_+ = \max\{0, s\}$ and $\chi_{[t_1,t_2]}$ is the characteristic function of $[t_1, t_2]$ $(0 \le t_1 < t_2 \le T)$. Let

$$I_k(t) = \int_\Omega (u_\varepsilon - k)_+^{m_1+1}dx; \tag{2.26}$$

then $I_k(t)$ is absolutely continuous on $[0, T]$. Denote by σ the point where $I_k(t)$ takes its maximum. Assume that $\sigma > 0$, for a sufficient small positive constant ϵ. Taking $t_1 = \sigma - \epsilon$, $t_2 = \sigma$ in (2.25), we obtain

$$\begin{aligned}
&\frac{1}{(m_1+1)\epsilon}\int_{\sigma-\epsilon}^{\sigma}\int_\Omega\frac{\partial(u_\varepsilon-k)_+^{m_1+1}}{\partial t}dx\,dt+\frac{1}{\epsilon}\int_{\sigma-\epsilon}^{\sigma}\int_\Omega\left|\nabla(u_\varepsilon-k)_+^{m_1}\right|^2dx\,dt\\
&\quad\le\frac{1}{\epsilon}\int_{\sigma-\epsilon}^{\sigma}\int_\Omega u_\varepsilon^{\alpha+m_1}(a+cv_\varepsilon)dx\,dt.
\end{aligned}\tag{2.27}$$

From

$$\int_{\sigma-\epsilon}^{\sigma}\int_\Omega\frac{\partial(u_\varepsilon-k)_+^{m_1+1}}{\partial t}dx\,dt = I_k(\sigma)-I_k(\sigma-\epsilon)\ge 0, \tag{2.28}$$

we have that

$$\frac{1}{\epsilon}\int_{\sigma-\epsilon}^{\sigma}\int_\Omega\left|\nabla(u_\varepsilon-k)_+^{m_1}\right|^2dx\,dt\le\frac{1}{\epsilon}\int_{\sigma-\epsilon}^{\sigma}\int_\Omega u_\varepsilon^{\alpha+m_1}(a+cv_\varepsilon)dx\,dt. \tag{2.29}$$

to Contents, page 1

Letting $\epsilon \to 0^+$, we have that

$$\int_\Omega \left|\nabla (u_\varepsilon(x,\sigma)-k)_+^{m_1}\right|^2 dx \le \int_\Omega u_\varepsilon^{\alpha+m_1}(x,\sigma)(a+cv_\varepsilon(x,\sigma))dx. \tag{2.30}$$

Denote $A_k(t) = \{x : u_\varepsilon(x,t) > k\}$ and $\mu_k = \sup_{t\in(0,T)}|A_k(t)|$; then

$$\int_{A_k(\sigma)} \left|\nabla (u_\varepsilon-k)_+^{m_1}\right|^2 dx \le \int_{A_k(\sigma)} u_\varepsilon^{\alpha+m_1}(a+cv_\varepsilon)dx. \tag{2.31}$$

By Sobolev's theorem,

$$\left(\int_{A_k(\sigma)} ((u_\varepsilon-k)_+^{m_1})^p dx\right)^{1/p} \le C\left(\int_{A_k(\sigma)} \left|\nabla (u_\varepsilon-k)_+^{m_1}\right|^2 dx\right)^{1/2}, \tag{2.32}$$

with

$$2 < p < \begin{cases} +\infty, & N \le 2, \\ \dfrac{2N}{N-2}, & N > 2, \end{cases} \tag{2.33}$$

we obtain

$$\begin{aligned}
\left(\int_{A_k(\sigma)} ((u_\varepsilon-k)_+^{m_1})^p dx\right)^{2/p} &\le C\int_{A_k(\sigma)} \left|\nabla (u_\varepsilon-k)_+^{m_1}\right|^2 dx \\
&\le C\int_{A_k(\sigma)} u_\varepsilon^{\alpha+m_1}(a+v_\varepsilon)dx \\
&\le C\left(\int_{A_k(\sigma)} u_\varepsilon^r dx\right)^{(m_1+\alpha)/r}\left(\int_{A_k(\sigma)} (a+v_\varepsilon)^{r/(r-m_1-\alpha)} dx\right)^{(r-m_1-\alpha)/r} \\
&\le C\left(\int_{A_k(\sigma)} (a+v_\varepsilon)^{r/(r-m_1-\alpha)} dx\right)^{(r-m_1-\alpha)/r} \\
&\le C\left(\int_{A_k(\sigma)} (a+v_\varepsilon)^s dx\right)^{1/s} |A_k(\sigma)|^{(s(r-m_1-\alpha)-r)/sr} \\
&\le C\mu_k^{(s(r-m_1-\alpha)-r)/sr},
\end{aligned} \tag{2.34}$$

to Contents, page 1

where $r > p(m_1 + \alpha)/(p - 2)$, $s > pr/(p(r - m_1 - \alpha) - 2r)$ and C denotes various positive constants independent of ε. By Hölder's inequality, it yields

$$\begin{aligned} I_k(\sigma) &= \int_{\Omega} (u_\varepsilon - k)_+^{m_1+1} dx = \int_{A_k(\sigma)} (u_\varepsilon - k)_+^{m_1+1} dx \\ &\le \left(\int_{A_k(\sigma)} (u_\varepsilon - k)_+^{m_1 p} dx \right)^{(m_1+1)/m_1 p} \mu_k^{1-(m_1+1)/m_1 p} \\ &\le C \mu_k^{1+[sp(r-m_1-\alpha)-pr-2sr](m_1+1)/2psrm_1}. \end{aligned} \tag{2.35}$$

Then

$$I_k(t) \le I_k(\sigma) \le C \mu_k^{1+[sp(r-m_1-\alpha)-pr-2sr](m_1+1)/2psrm_1}, \qquad t \in [0, T]. \tag{2.36}$$

On the other hand, for any $h > k$ and $t \in [0, T]$, we have that

$$I_k(t) \ge \int_{A_k(t)} (u_\varepsilon - k)_+^{m_1+1} dx \ge (h - k)^{m_1+1} |A_h(t)|. \tag{2.37}$$

Combined with (2.35), it yields

$$(h - k)^{m_1+1} \mu_h \le C \mu_k^{1+[sp(r-m_1-\alpha)-pr-2sr](m_1+1)/2psrm_1}, \tag{2.38}$$

that is,

$$\mu_h \le \frac{C}{(h - k)^{m_1+1}} \mu_k^{1+[sp(r-m_1-\alpha)-pr-2sr](m_1+1)/2psrm_1}. \tag{2.39}$$

It is easy to see that

$$\gamma = 1 + \frac{[sp(r - m_1 - \alpha) - pr - 2sr](m_1 + 1)}{2psrm_1} > 1. \tag{2.40}$$

Then by the De Giorgi iteration lemma [22], we have that

$$\mu_{l+d} = \sup |A_{l+d}(t)| = 0, \tag{2.41}$$

where $d = C^{1/(m_1+1)} \mu_l^{(\gamma-1)/(m_1+1)} 2^{\gamma/(\gamma-1)}$. That is,

$$u_\varepsilon \le l + d \quad \text{a.e. in } Q_T. \tag{2.42}$$

It is the same for the second inequality of (2.24). The proof is completed. □

to Contents, page 1

Lemma 2.4. *The solution* $(u_\varepsilon, v_\varepsilon)$ *of* (2.1)–(2.4) *satisfies the following:*

$$\iint_{Q_T} \left| \frac{\partial u_\varepsilon^{m_1}}{\partial t} \right|^2 dx\,dt \le C, \qquad \iint_{Q_T} \left| \frac{\partial v_\varepsilon^{m_2}}{\partial t} \right|^2 dx\,dt \le C, \tag{2.43}$$

where C *is a positive constant independent of* ε.

Proof. Multiplying (2.1) by $(\partial/\partial t)u_\varepsilon^{m_1}$ and integrating over Ω, by (2.3), (2.4) and Young's inequality we have that

$$\begin{aligned}
&\frac{4m_1}{(m_1+1)^2} \iint_{Q_T} \left| \frac{\partial}{\partial t} u_\varepsilon^{(m_1+1)/2} \right|^2 dx\,dt \\
&\quad = \iint_{Q_T} \frac{\partial u_\varepsilon}{\partial t} \frac{\partial u_\varepsilon^{m_1}}{\partial t} dx\,dt \\
&\quad = \frac{1}{2}\int_\Omega |\nabla u_\varepsilon^{m_1}(x,0)|^2 dx - \frac{1}{2}\int_\Omega |\nabla u_\varepsilon^{m_1}(x,T)|^2 dx \\
&\qquad + \iint_{Q_T} m_1 u_\varepsilon^{\alpha+m_1-1}(a - bu_\varepsilon + cv_\varepsilon)\frac{\partial u_\varepsilon}{\partial t} dx\,dt \\
&\quad = \frac{1}{2}\int_\Omega |\nabla u_\varepsilon^{m_1}(x,0)|^2 dx - \frac{1}{2}\int_\Omega |\nabla u_\varepsilon^{m_1}(x,T)|^2 dx \\
&\qquad + \iint_{Q_T} \frac{2m_1}{m_1+1} u_\varepsilon^{(2\alpha+m_1-1)/2}(a - bu_\varepsilon + cv_\varepsilon)\frac{\partial u_\varepsilon^{(m_1+1)/2}}{\partial t} dx\,dt \\
&\quad \le \frac{1}{2}\int_\Omega |\nabla u_\varepsilon^{m_1}(x,0)|^2 dx + 2m_1 \iint_{Q_T} u_\varepsilon^{2\alpha+m_1-1}(a - bu_\varepsilon + cv_\varepsilon)^2 dx\,dt \\
&\qquad + \frac{2m_1}{(m_1+1)^2} \iint_{Q_T} \left| \frac{\partial}{\partial t} u_\varepsilon^{(m_1+1)/2} \right|^2 dx\,dt,
\end{aligned} \tag{2.44}$$

which together with the bound of $a, b, c, u_\varepsilon, v_\varepsilon$ shows that

$$\iint_{Q_T} \left| \frac{\partial u_\varepsilon^{(m_1+1)/2}}{\partial t} \right|^2 dx\,dt \le C, \tag{2.45}$$

where C is a positive constant independent of ε. Noticing the bound of u_ε, we have that

$$\iint_{Q_T} \left| \frac{\partial u_\varepsilon^{m_1}}{\partial t} \right|^2 dx\,dt = \frac{4m_1^2}{(m_1+1)^2} \iint_{Q_T} u_\varepsilon^{m_1-1} \left| \frac{\partial}{\partial t} u_\varepsilon^{(m_1+1)/2} \right|^2 dx\,dt \le C. \tag{2.46}$$

It is the same for the second inequality. The proof is completed. □

From the above estimates of $u_\varepsilon, v_\varepsilon$, we have the following results.

to Contents, page 1

Theorem 2.5. *The problem* (1.1)–(1.4) *admits a generalized solution.*

Proof. By Lemmas 2.2, 2.3, and 2.4, we can see that there exist subsequences of $\{u_\varepsilon\}, \{v_\varepsilon\}$ (denoted by themselves for simplicity) and functions u, v such that

$$\begin{gathered} u_\varepsilon \longrightarrow u, \quad v_\varepsilon \longrightarrow v, \quad \text{a.e in } Q_T, \\ \frac{\partial u_\varepsilon^{m_1}}{\partial t} \longrightarrow \frac{\partial u^{m_1}}{\partial t}, \quad \frac{\partial v_\varepsilon^{m_2}}{\partial t} \longrightarrow \frac{\partial v^{m_2}}{\partial t}, \quad \text{weakly in } L^2(Q_T), \\ \nabla u_\varepsilon^{m_1} \longrightarrow \nabla u^{m_1}, \quad \nabla v_\varepsilon^{m_2} \longrightarrow \nabla v^{m_2}, \quad \text{weakly in } L^2(Q_T), \end{gathered} \tag{2.47}$$

as $\varepsilon \to 0$. Then a rather standard argument as [23] shows that (u, v) is a generalized solution of (1.1)–(1.4) in the sense of Definition 1.1. □

In order to prove that the generalized solution of (1.1)–(1.4) is uniformly bounded, we need the following comparison principle.

Lemma 2.6. *Let* $(\underline{u}, \underline{v})$ *be a subsolution of the problem* (1.1)–(1.4) *with the initial value* $(\underline{u}_0, \underline{v}_0)$ *and* $(\overline{u}, \overline{v})$ *a supersolution with a positive lower bound of the problem* (1.1)–(1.4) *with the initial value* $(\overline{u}_0, \overline{v}_0)$. *If* $\underline{u}_0 \le \overline{v}_0$, $\underline{u}_0 \le \overline{v}_0$, *then* $\underline{u}(x,t) \le \overline{u}(x,t)$, $\underline{v}(x,t) \le \overline{v}(x,t)$ *on* Q_T.

Proof. Without loss of generality, we might assume that $\|\underline{u}(x,t)\|_{L^\infty(Q_T)}$, $\|\overline{u}(x,t)\|_{L^\infty(Q_T)}$, $\|\underline{v}(x,t)\|_{L^\infty(Q_T)}$, $\|\overline{v}(x,t)\|_{L^\infty(Q_T)} \le M$, where M is a positive constant. By the definitions of subsolution and supersolution, we have that

$$\begin{aligned} &\int_0^t \int_\Omega -\underline{u}\frac{\partial \varphi}{\partial t} + \nabla \underline{u}^{m_1} \nabla \varphi dx\, d\tau + \int_\Omega \underline{u}(x,t)\varphi(x,t)dx - \int_\Omega \underline{u}_0(x)\varphi(x,0)dx \\ &\quad \le \int_0^t \int_\Omega \underline{u}^\alpha (a - b\underline{u} + c\underline{v})\varphi dx\, d\tau, \\ &\int_0^t \int_\Omega -\overline{u}\frac{\partial \varphi}{\partial t} + \nabla \overline{u}^{m_1} \nabla \varphi dx\, d\tau + \int_\Omega \overline{u}(x,t)\varphi(x,t)dx - \int_\Omega \overline{v}_0(x)\varphi(x,0)dx \\ &\quad \ge \int_0^t \int_\Omega \overline{u}^\alpha (a - b\overline{u} + c\overline{v})\varphi dx\, d\tau. \end{aligned} \tag{2.48}$$

Take the test function as

$$\varphi(x,t) = H_\varepsilon\left(\underline{u}^{m_1}(x,t) - \overline{u}^{m_1}(x,t)\right), \tag{2.49}$$

where $H_\varepsilon(s)$ is a monotone increasing smooth approximation of the function $H(s)$ defined as follows:

$$H(s) = \begin{cases} 1, & s > 0, \\ 0, & \text{otherwise}. \end{cases} \tag{2.50}$$

to Contents, page 1

It is easy to see that $H'_\varepsilon(s) \to \delta(s)$ as $\varepsilon \to 0$. Since $\partial \underline{u}^{m_1}/\partial t$, $\partial \overline{u}^{m_1}/\partial t \in L^2(Q_T)$, the test function $\varphi(x,t)$ is suitable. By the positivity of a, b, c we have that

$$\begin{aligned}\int_\Omega (\underline{u}-\overline{u}) H_\varepsilon(\underline{u}^{m_1}-\overline{u}^{m_1})dx - \int_0^t \int_\Omega (\underline{u}-\overline{u}) \frac{\partial H_\varepsilon(\underline{u}^{m_1}-\overline{u}^{m_1})}{\partial t} dx\, d\tau \\ + \int_0^t \int_\Omega H'_\varepsilon(\underline{u}^{m_1}-\overline{u}^{m_1}) \left|\nabla(\underline{u}^{m_1}-\overline{u}^{m_1})\right|^2 dx\, d\tau \\ \leq \int_0^t \int_\Omega a(\underline{u}^\alpha - \overline{u}^\alpha) H_\varepsilon(\underline{u}^{m_1}-\overline{u}^{m_1}) + c(\underline{u}^\alpha \underline{v} - \overline{u}^\alpha \overline{v}) H_\varepsilon(\underline{u}^{m_1}-\overline{u}^{m_1}) dx\, d\tau,\end{aligned} \tag{2.51}$$

where C is a positive constant depending on $\|a(x,t)\|_{C(Q_t)}$, $\|c(x,t)\|_{C(Q_t)}$. Letting $\varepsilon \to 0$ and noticing that

$$\int_0^t \int_\Omega H'_\varepsilon(\underline{u}^{m_1}-\overline{u}^{m_1}) \left|\nabla(\underline{u}^{m}-\overline{u}^{m})\right|^2 dx\, d\tau \geq 0, \tag{2.52}$$

we arrive at

$$\int_\Omega \left[\underline{u}(x,t) - \overline{u}(x,t)\right]_+ dx \leq C \int_0^t \int_\Omega (\underline{u}^\alpha - \overline{u}^\alpha)_+ + \underline{v}(\underline{u}^\alpha - \overline{u}^\alpha)_+ + \overline{u}^\alpha(\underline{v}-\overline{v})_+ dx\, d\tau. \tag{2.53}$$

Let $(\overline{u}, \overline{v})$ be a supsolution with a positive lower bound σ. Noticing that

$$\begin{aligned}(x^\alpha - y^\alpha)_+ &\leq C(\alpha)(x-y)_+, \quad \text{for } \alpha \geq 1, \\ (x^\alpha - y^\alpha)_+ &\leq x^{\alpha-1}(x-y)_+ \leq y^{\alpha-1}(x-y)_+, \quad \text{for } \alpha < 1,\end{aligned} \tag{2.54}$$

with $x, y > 0$, we have that

$$\int_0^t \int_\Omega (\underline{u}^\alpha - \overline{u}^\alpha)_+ + \underline{v}(\underline{u}^\alpha - \overline{u}^\alpha)_+ + \overline{u}^\alpha(\underline{v}-\overline{v})_+ dx\, d\tau \leq C \int_0^t \int_\Omega (\underline{u}-\overline{u})_+ + (\underline{v}-\overline{v})_+ dx\, d\tau, \tag{2.55}$$

where C is a positive constant depending upon α, σ, M.

Similarly, we also have that

$$\int_\Omega \left[\underline{v}(x,t) - \overline{v}(x,t)\right]_+ dx \leq C \int_0^t \int_\Omega (\underline{u}-\overline{u})_+ + (\underline{v}-\overline{v})_+ dx\, d\tau. \tag{2.56}$$

to Contents, page 1

Combining the above two inequalities, we obtain

$$\int_{\Omega} [\underline{u}(x,t) - \overline{u}(x,t)]_+ + [\underline{v}(x,t) - \overline{v}(x,t)]_+ dx \le C \int_0^t \int_{\Omega} (\underline{u} - \overline{u})_+ + (\underline{v} - \overline{v})_+ dx\, d\tau. \tag{2.57}$$

By Gronwall's lemma, we see that $\underline{u} \le \overline{u}$, $\underline{v} \le \overline{v}$. The proof is completed. □

Corollary 2.7. *If $b_l f_l > c_M e_M$, then the problem* (1.1)–(1.4) *admits at most one global solution which is uniformly bounded in $\overline{\Omega} \times [0, \infty)$.*

Proof. The uniqueness comes from the comparison principle immediately. In order to prove that the solution is global, we just need to construct a bounded positive supersolution of (1.1)–(1.4).

Let $\rho_1 = (a_M f_l + d_M c_M)/(b_l f_l - c_M e_M)$ and $\rho_2 = (a_M e_M + d_M b_l)/(b_l f_l - c_M e_M)$, since $b_l f_l > c_M e_M$; then $\rho_1, \rho_2 > 0$ and satisfy

$$a_M - b_l \rho_1 + c_M \rho_2 = 0, \qquad d_M + e_M \rho_1 - f_l \rho_2 = 0. \tag{2.58}$$

Let $(\overline{u}, \overline{v}) = (\eta\rho_1, \eta\rho_2)$, where $\eta > 1$ is a constant such that $(u_0, v_0) \le (\eta\rho_1, \eta\rho_2)$; then we have that

$$\overline{u}_t - \Delta \overline{u}^{m_1} = 0 \ge \overline{u}^{\alpha}(a - b\overline{u} + c\overline{v}), \qquad \overline{v}_t - \Delta \overline{v}^{m_2} = 0 \ge \overline{v}^{\beta}(d + e\overline{u} - f\overline{v}). \tag{2.59}$$

That is, $(\overline{u}, \overline{v}) = (\eta\rho_1, \eta\rho_2)$ is a positive supersolution of (1.1)–(1.4). Since $\overline{u}, \overline{v}$ are global and uniformly bounded, so are u and v. □

3. Periodic Solutions

In order to establish the existence of the nontrivial nonnegative periodic solutions of the problem (1.1)–(1.3), we need the following lemmas. Firstly, we construct a pair of T-periodic supersolution and T-periodic subsolution as follows.

Lemma 3.1. *In case of $b_l f_l > c_M e_M$, there exists a pair of T-periodic supersolution and T-periodic subsolution of the problem* (1.1)–(1.3).

Proof. We first construct a T-periodic subsolution of (1.1)–(1.3). Let λ be the first eigenvalue and ϕ be the uniqueness solution of the following elliptic problem:

$$-\Delta\phi = \lambda\phi, \quad x \in \Omega, \quad \phi = 0, \quad x \in \partial\Omega; \tag{3.1}$$

then we have that

$$\lambda > 0, \quad \phi(x) > 0 \quad \text{in } \Omega, \quad |\nabla\phi| > 0 \quad \text{on } \partial\Omega, \quad M = \max_{x \in \overline{\Omega}} \phi(x) < \infty. \tag{3.2}$$

to Contents, page 1

Let

$$(\underline{u},\underline{v}) = \left(\varepsilon\phi^{2/m_1}(x), \varepsilon\phi^{2/m_2}(x)\right), \tag{3.3}$$

where $\varepsilon > 0$ is a small constant to be determined. We will show that $(\underline{u},\underline{v})$ is a (time independent, hence T-periodic) subsolution of (1.1)–(1.3).

Taking the nonnegative function $\varphi_1(x,t) \in C^1(\overline{Q}_T)$ as the test function, we have that

$$\begin{aligned}
&\iint_{Q_T}\left(\underline{u}\frac{\partial\varphi_1}{\partial t} + \Delta\underline{u}^{m_1}\varphi_1 + \underline{u}^{\alpha}(a - b\underline{u} + c\underline{v})\varphi_1\right)dx\,dt \\
&\qquad + \int_{\Omega}\underline{u}(x,0)\varphi_1(x,0) - \underline{u}(x,T)\varphi_1(x,T)dx \\
&\quad = \iint_{Q_T}(\underline{u}^{\alpha}(a - b\underline{u} + c\underline{v}) + \Delta\underline{u}^{m_1})\varphi_1 dx\,dt \\
&\quad = \iint_{Q_T}\underline{u}^{\alpha}(a - b\underline{u} + c\underline{v})\varphi_1 dx\,dt - \iint_{Q_T}\nabla\underline{u}^{m_1}\nabla\varphi_1 dx\,dt \\
&\quad = \iint_{Q_T}\underline{u}^{\alpha}(a - b\underline{u} + c\underline{v})\varphi_1 dx\,dt - 2\varepsilon^{m_1}\iint_{Q_T}\phi\nabla\phi\cdot\nabla\varphi_1 dx\,dt \\
&\quad = \iint_{Q_T}\underline{u}^{\alpha}(a - b\underline{u} + c\underline{v})\varphi_1 dx\,dt - 2\varepsilon^{m_1}\iint_{Q_T}\nabla\phi\nabla(\phi\varphi_1) - |\nabla\phi|^2\varphi_1 dx\,dt \\
&\quad = \iint_{Q_T}\underline{u}^{\alpha}(a - b\underline{u} + c\underline{v})\varphi_1 dx\,dt - 2\varepsilon^{m_1}\iint_{Q_T} -\operatorname{div}(\nabla\phi)\phi\varphi_1 - |\nabla\phi|^2\varphi_1 dx\,dt \\
&\quad = \iint_{Q_T}\underline{u}^{\alpha}(a - b\underline{u} + c\underline{v})\varphi_1 dx\,dt - 2\varepsilon^{m_1}\iint_{Q_T}\left(\lambda\phi^2 - |\nabla\phi|^2\right)\varphi_1 dx\,dt.
\end{aligned} \tag{3.4}$$

Similarly, for any nonnegative test function $\varphi_2(x,t) \in C^1(\overline{Q}_T)$, we have that

$$\begin{aligned}
&\iint_{Q_T}\left(\underline{v}\frac{\partial\varphi_2}{\partial t} + \Delta\underline{v}^{m_2}\varphi_2 + \underline{v}^{\beta}(d + e\underline{u} - f\underline{v})\varphi_2\right)dx\,dt + \int_{\Omega}\underline{v}(x,0)\varphi_2(x,0) - \underline{v}(x,T)\varphi_2(x,T)dx \\
&\quad = \iint_{Q_T}\underline{v}^{\beta}(d + e\underline{u} - f\underline{v})\varphi_2 dx\,dt - 2\varepsilon^{m_2}\iint_{Q_T}\left(\lambda\phi^2 - |\nabla\phi|^2\right)\varphi_2 dx\,dt.
\end{aligned} \tag{3.5}$$

We just need to prove the nonnegativity of the right-hand side of (3.4) and (3.5).

Since $\phi_1 = \phi_2 = 0$, $|\nabla\phi_1|, |\nabla\phi_2| > 0$ on $\partial\Omega$, then there exists $\delta > 0$ such that

$$\lambda\phi^2 - |\nabla\phi|^2 \le 0, \qquad x \in \overline{\Omega}_{\delta}, \tag{3.6}$$

to Contents, page 1

where $\overline{\Omega}_\delta = \{x \in \Omega \mid \operatorname{dist}(x, \partial\Omega) \le \delta\}$. Choosing

$$\varepsilon \le \min\left\{ \frac{a_l}{b_M M^{2/m_1}}, \frac{d_l}{f_M M^{2/m_2}} \right\}, \tag{3.7}$$

then we have that

$$\begin{aligned} 2\varepsilon^{m_1} \int_0^T \int_{\Omega_\delta} \left(\lambda\phi^2 - |\nabla\phi|^2\right)\varphi_1 dx\, dt \le 0 \le \int_0^T \int_{\Omega_\delta} \underline{u}^{\alpha}(a - b\underline{u} + c\underline{v})\varphi_1 dx\, dt, \\ 2\varepsilon^{m_2} \int_0^T \int_{\Omega_\delta} \left(\lambda\phi^2 - |\nabla\phi|^2\right)\varphi_2 dx\, dt \le 0 \le \int_0^T \int_{\Omega_\delta} \underline{v}^{\beta}(d + e\underline{u} - f\underline{v})\varphi_2 dx\, dt, \end{aligned} \tag{3.8}$$

which shows that $(\underline{u}, \underline{v})$ is a positive (time independent, hence T-periodic) subsolution of (1.1)–(1.3) on $\overline{\Omega}_\delta \times (0, T)$.

Moreover, we can see that, for some $\sigma > 0$,

$$\phi(x) \ge \sigma > 0, \quad x \in \Omega \setminus \overline{\Omega}_\delta. \tag{3.9}$$

Choosing

$$\varepsilon \le \min\left\{ \frac{a_l}{2b_M M^{2/m_1}}, \left(\frac{a_l}{4\lambda M^{2(m_1-\alpha)/m_1}}\right)^{1/(m_1-\alpha)}, \frac{d_l}{2f_M M^{2/m_2}}, \left(\frac{d_l}{4\lambda M^{2(m_2-\beta)/m_2}}\right)^{1/(m_2-\beta)} \right\}, \tag{3.10}$$

then

$$\begin{aligned} \varepsilon^{\alpha}\phi^{2\alpha/m_1} a - b\varepsilon^{\alpha+1}\phi^{2(\alpha+1)/m_1} + c\varepsilon^{\alpha}\phi^{2\alpha/m_1}\varepsilon\phi^{2/m_2} - 2\varepsilon^{m_1}\lambda\phi^2 \ge 0, \\ \varepsilon^{\beta}\phi^{2\beta/m_2} d + e\varepsilon\phi^{2/m_1}\varepsilon^{\beta}\phi^{2\beta/m_2} - f\varepsilon^{\beta+1}\phi^{2(\beta+1)/m_2} - 2\varepsilon^{m_2}\lambda\phi^2 \ge 0 \end{aligned} \tag{3.11}$$

on Q_T, that is

$$\begin{aligned} \iint_{Q_T} \underline{u}^{\alpha}(a - b\underline{u} + c\underline{v})\varphi_1 dx\, dt - 2\varepsilon^{m_1} \iint_{Q_T} \left(\lambda\phi^2 - |\nabla\phi|^2\right)\varphi_1 dx\, dt \ge 0, \\ \iint_{Q_T} \underline{v}^{\beta}(d + e\underline{u} - f\underline{v})\varphi_2 dx\, dt - 2\varepsilon^{m_2} \iint_{Q_T} \left(\lambda\phi^2 - |\nabla\phi|^2\right)\varphi_2 dx\, dt \ge 0. \end{aligned} \tag{3.12}$$

These relations show that $(\underline{u}, \underline{v}) = (\varepsilon\phi_1^{2/m_1}(x), \varepsilon\phi_2^{2/m_2}(x))$ is a positive (time independent, hence T-periodic) subsolution of (1.1)–(1.3).

Letting $(\overline{u}, \overline{v}) = (\eta\rho_1, \eta\rho_2)$, where η, ρ_1, ρ_2 are taken as those in Corollary 2.7, it is easy to see that $(\overline{u}, \overline{v})$ is a positive (time independent, hence T-periodic) subsolution of (1.1)–(1.3).

Obviously, we may assume that $\underline{u}(x,t) \le \overline{u}(x,t)$, $\underline{v}(x,t) \le \overline{v}(x,t)$ by changing η, ε appropriately. □

to Contents, page 1

Lemma 3.2 (see [24, 25]). *Let u be the solution of the following Dirichlet boundary value problem*

$$\frac{\partial u}{\partial t} = \Delta u^m + f(x,t), \quad (x,t) \in \Omega \times (0,T),$$
$$u(x,t) = 0, \quad (x,t) \in \partial\Omega \times (0,T), \tag{3.13}$$

where $f \in L^\infty(\Omega \times (0,T))$; then there exist positive constants K and $\alpha \in (0,1)$ depending only upon $\tau \in (0,T)$ and $\|f\|_{L^\infty(\Omega\times(0,T))}$, such that, for any $(x_i,t_i) \in \Omega \times [\tau,T]$ $(i = 1,2)$,

$$|u(x_1,t_1) - u(x_2,t_2)| \le K\left(|x_1 - x_2|^\alpha + |t_1 - t_2|^{\alpha/2}\right). \tag{3.14}$$

Lemma 3.3 (see [26]). *Define a Poincaré mapping*

$$P_t : L^\infty(\Omega) \times L^\infty(\Omega) \longrightarrow L^\infty(\Omega) \times L^\infty(\Omega),$$
$$P_t(u_0(x), v_0(x)) := (u(x,t), v(x,t)) \quad (t > 0), \tag{3.15}$$

where $(u(x,t), v(x,t))$ is the solution of (1.1)–(1.4) with initial value $(u_0(x), v_0(x))$. According to Lemmas 2.6 and 3.2 and Theorem 2.5, the map P_t has the following properties:

(i) *P_t is defined for any $t > 0$ and order preserving;*

(ii) *P_t is order preserving;*

(iii) *P_t is compact.*

Observe that the operator P_T is the classical Poincaré map and thus a fixed point of the Poincaré map gives a T-periodic solution setting. This will be made by the following iteration procedure.

Theorem 3.4. *Assume that $b_l f_l > c_M e_M$ and there exists a pair of nontrivial nonnegative T-periodic subsolution $(\underline{u}(x,t), \underline{v}(x,t))$ and T-periodic supersolution $(\overline{u}(x,t), \overline{v}(x,t))$ of the problem (1.1)–(1.3) with $\underline{u}(x,0) \le \overline{u}(x,0)$; then the problem (1.1)–(1.3) admits a pair of nontrivial nonnegative periodic solutions $(u_*(x,t), v_*(x,t))$, $(u^*(x,t), v^*(x,t))$ such that*

$$\underline{u}(x,t) \le u_*(x,t) \le u^*(x,t) \le \overline{u}(x,t), \quad \underline{v}(x,t) \le v_*(x,t) \le v^*(x,t) \le \overline{v}(x,t), \quad \text{in } Q_T. \tag{3.16}$$

Proof. Taking $\overline{u}(x,t), \underline{u}(x,t)$ as those in Lemma 3.1 and choosing suitable $B(x_0,\delta), B(x_0,\delta'), \Omega'$, k_1, k_2, and K, we can obtain $\underline{u}(x,0) \le \overline{u}(x,0)$. By Lemma 2.6, we have that $P_T(\underline{u}(\cdot,0)) \ge \underline{u}(\cdot,T)$. Hence by Definition 1.2 we get $P_T(\underline{u}(\cdot,0)) \ge \underline{u}(\cdot,0)$, which implies $P_{(k+1)T}(\underline{u}(\cdot,0)) \ge P_{kT}(\underline{u}(\cdot,0))$ for any $k \in \mathbb{N}$. Similarly we have that $P_T(\overline{u}(\cdot,0)) \le \overline{u}(\cdot,T) \le \overline{u}(\cdot,0)$, and hence $P_{(k+1)T}(\overline{u}(\cdot,0)) \le P_{kT}(\overline{u}(\cdot,0))$ for any $k \in \mathbb{N}$. By Lemma 2.6, we have that $P_{kT}(\underline{u}(\cdot,0)) \le P_{kT}(\overline{u}(\cdot,0))$ for any $k \in \mathbb{N}$. Then

$$u_*(x,0) = \lim_{k\to\infty} P_{kT}(\underline{u}(x,0)), \qquad u^*(x,0) = \lim_{k\to\infty} P_{kT}(\overline{u}(x,0)) \tag{3.17}$$

exist for almost every $x \in \Omega$. Since the operator P_T is compact (see Lemma 3.3), the above limits exist in $L^\infty(\Omega)$, too. Moreover, both $u_*(x,0)$ and $u^*(x,0)$ are fixed points of P_T. With

to Contents, page 1

the similar method as [26], it is easy to show that the even extension of the function $u_*(x,t)$, which is the solution of the problem (1.1)–(1.4) with the initial value $u_*(x,0)$, is indeed a nontrivial nonnegative periodic solution of the problem (1.1)–(1.3). It is the same for the existence of $u^*(x,t)$. Furthermore, by Lemma 2.6, we obtain (3.16) immediately, and thus we complete the proof. □

Furthermore, by De Giorgi iteration technique, we can also establish a prior upper bound of all nonnegative periodic solutions of (1.1)–(1.3). Then with a similar method as [18], we have the following remark which shows the existence and attractivity of the maximal periodic solution.

Remark 3.5. If $b_l f_l > c_M e_M$, the problem (1.1)–(1.3) admits a maximal periodic solution (U,V). Moreover, if (u,v) is the solution of the initial boundary value problem (1.1)–(1.4) with nonnegative initial value (u_0,v_0), then, for any $\varepsilon > 0$, there exists t depending on u_0, v_0, and ε, such that

$$0 \le u \le U + \varepsilon, \quad 0 \le v \le V + \varepsilon, \quad \text{for } x \in \Omega,\ t \ge t. \tag{3.18}$$

Acknowledgments

This work was supported by NSFC (10801061), the Fundamental Research Funds for the Central Universities (Grant no. HIT. NSRIF. 2009049), Natural Sciences Foundation of Heilongjiang Province (Grant no. A200909), and also the 985 project of Harbin Institute of Technology.

References

[1] J. A. Cui, "Global asymptotic stability in n-species cooperative system with time delays," *Systems Science and Mathematical Sciences*, vol. 7, no. 1, pp. 45–48, 1994.

[2] D. Hu and Z. Zhang, "Four positive periodic solutions to a Lotka-Volterra cooperative system with harvesting terms," *Nonlinear Analysis: Real World Applications*, vol. 11, no. 2, pp. 1115–1121, 2010.

[3] K. I. Kim and Z. Lin, "A degenerate parabolic system with self-diffusion for a mutualistic model in ecology," *Nonlinear Analysis: Real World Applications*, vol. 7, no. 4, pp. 597–609, 2006.

[4] Y. Lou, T. Nagylaki, and W.-M. Ni, "On diffusion-induced blowups in a mutualistic model," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 45, no. 3, pp. 329–342, 2001.

[5] Z. Lin, J. Liu, and M. Pedersen, "Periodicity and blowup in a two-species cooperating model," *Nonlinear Analysis: Real World Applications*, vol. 12, no. 1, pp. 479–486, 2011.

[6] C. V. Pao, *Nonlinear Parabolic and Elliptic Equations*, Plenum Press, New York, NY, USA, 1992.

[7] Z. Y. Lu and Y. Takeuchi, "Permanence and global stability for cooperative Lotka-Volterra diffusion systems," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 19, no. 10, pp. 963–975, 1992.

[8] J. Sun, B. Wu, and D. Zhang, "Asymptotic behavior of solutions of a periodic diffusion equation," *Journal of Inequalities and Applications*, vol. 2010, Article ID 597569, 12 pages, 2010.

[9] S. Ahmad and A. C. Lazer, "Asymptotic behaviour of solutions of periodic competition diffusion system," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 13, no. 3, pp. 263–284, 1989.

[10] A. Tineo, "Existence of global coexistence state for periodic competition diffusion systems," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 19, no. 4, pp. 335–344, 1992.

[11] K. Gopalsamy, "Global asymptotic stability in a periodic Lotka-Volterra system," *Journal of the Australian Mathematical Society. Series B*, vol. 27, no. 1, pp. 66–72, 1985.

[12] J. López-Gómez, "Positive periodic solutions of Lotka-Volterra reaction-diffusion systems," *Differential and Integral Equations*, vol. 5, no. 1, pp. 55–72, 1992.

[13] J. J. Morgan and S. L. Hollis, "The existence of periodic solutions to reaction-diffusion systems with periodic data," *SIAM Journal on Mathematical Analysis*, vol. 26, no. 5, pp. 1225–1232, 1995.

to Contents, page 1

[14] C. V. Pao, "Periodic solutions of parabolic systems with nonlinear boundary conditions," *Journal of Mathematical Analysis and Applications*, vol. 234, no. 2, pp. 695–716, 1999.

[15] C. Tian and Z. Lin, "Periodic solutions of reaction diffusion systems in a half-space domain," *Nonlinear Analysis: Real World Applications*, vol. 9, no. 3, pp. 811–821, 2008.

[16] K. J. Brown and P. Hess, "Positive periodic solutions of predator-prey reaction-diffusion systems," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 16, no. 12, pp. 1147–1158, 1991.

[17] J.-B. Sun, C.-H. Jin, and Y.-Y. Ke, "Existence of non-trivial nonnegative periodic solutions for a nonlinear diffusion system," *Northeastern Mathematical Journal*, vol. 23, no. 2, pp. 167–175, 2007.

[18] J.-B. Sun, "Asymptotic bounds for solutions of a periodic reaction diffusion system," *Applied Mathematics E-Notes*, vol. 10, pp. 128–135, 2010.

[19] J. Yin and Y. Wang, "Asymptotic behaviour of solutions for nonlinear diffusion equation with periodic absorption," in *Partial Differential Equations and Their Applications (Wuhan, 1999)*, pp. 305–308, World Scientific, River Edge, NJ, USA, 1999.

[20] A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1964.

[21] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva, *Linear and Quasilinear Equations of Parabolic Type*, vol. 23 of *Translations of Mathematical Monographs*, American Mathematical Society, Providence, RI, USA, 1967.

[22] Z. Q. Wu, J. X. Yin, and C. P. Wang, *Introduction to Elliptic and Parabolic Equations*, Scientific Publications, Beijing, China, 2003.

[23] Z. Wu, J. Zhao, J. Yin, and H. Li, *Nonlinear Diffusion Equations*, World Scientific, River Edge, NJ, USA, 2001.

[24] E. DiBenedetto, "Continuity of weak solutions to a general porous medium equation," *Indiana University Mathematics Journal*, vol. 32, no. 1, pp. 83–118, 1983.

[25] M. M. Porzio and V. Vespri, "Hölder estimates for local solutions of some doubly nonlinear degenerate parabolic equations," *Journal of Differential Equations*, vol. 103, no. 1, pp. 146–178, 1993.

[26] P. Hess, M. A. Pozio, and A. Tesei, "Time periodic solutions for a class of degenerate parabolic problems," *Houston Journal of Mathematics*, vol. 21, no. 2, pp. 367–394, 1995.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 671967, 24 pages
doi:10.1155/2011/671967

Research Article

Asymptotic Behavior of Solutions of Delayed Difference Equations

J. Diblík[1,2] and I. Hlavičková[2]

[1] *Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, University of Technology, 602 00 Brno, Czech Republic*
[2] *Department of Mathematics, Faculty of Electrical Engineering and Communication, Brno University of Technology, 616 00 Brno, Czech Republic*

Correspondence should be addressed to J. Diblík, diblik@feec.vutbr.cz

Received 24 January 2011; Accepted 9 May 2011

Academic Editor: Miroslava Růžičková

This contribution is devoted to the investigation of the asymptotic behavior of delayed difference equations with an integer delay. We prove that under appropriate conditions there exists at least one solution with its graph staying in a prescribed domain. This is achieved by the application of a more general theorem which deals with systems of first-order difference equations. In the proof of this theorem we show that a good way is to connect two techniques—the so-called retract-type technique and Liapunov-type approach. In the end, we study a special class of delayed discrete equations and we show that there exists a positive and vanishing solution of such equations.

1. Introduction

Throughout this paper, we use the following notation: for an integer q, we define

$$\mathbb{Z}_q^\infty := \{q, q+1, \dots\}. \tag{1.1}$$

We investigate the asymptotic behavior for $n \to \infty$ of the solutions of the discrete delayed equation of the $(k+1)$-th order

$$\Delta v(n) = f(n, v(n), v(n-1), \dots, v(n-k)), \tag{1.2}$$

where n is the independent variable assuming values from the set $\mathbb{Z}_a^\infty$ with a fixed $a \in \mathbb{N}$. The number $k \in \mathbb{N}$, $k \geq 1$ is the fixed delay, $\Delta v(n) = v(n+1) - v(n)$, and $f : \mathbb{Z}_a^\infty \times \mathbb{R}^{k+1} \to \mathbb{R}$.

to Contents, page 1

A function $v : \mathbb{Z}_{a-k}^{\infty} \to \mathbb{R}$ is a solution of (1.2) if it satisfies (1.2) for every $n \in \mathbb{Z}_a^{\infty}$. We will study (1.2) together with $k+1$ initial conditions

$$v(a+s-k) = v^{a+s-k} \in \mathbb{R}, \quad s = 0, 1, \ldots, k. \tag{1.3}$$

Initial problem (1.2), (1.3) obviously has a unique solution, defined for every $n \in \mathbb{Z}_{a-k}^{\infty}$. If the function f is continuous with respect to its last $k+1$ arguments, then the solution of (1.2) continuously depends on initial conditions (1.3).

Now we give a general description of the problem solved in this paper.

Problem 1. Let $b, c : \mathbb{Z}_{a-k}^{\infty} \to \mathbb{R}$ be functions such that $b(n) < c(n)$ for every $n \in \mathbb{Z}_{a-k}^{\infty}$. The problem under consideration is to find sufficient conditions for the right-hand side of (1.2) that will guarantee the existence of a solution $v = v^*(n)$ of initial problem (1.2), (1.3) such that

$$b(n) < v^*(n) < c(n), \quad n \in \mathbb{Z}_{a-k}^{\infty}. \tag{1.4}$$

This problem can be solved with help of a result which is valid for systems of first-order difference equations and which will be presented in the next section. This is possible because the considered equation (1.2) can be rewritten as a system of $k+1$ first-order difference equations, similarly as a differential equation of a higher order can be transformed to a special system of first-order differential equations. Although the process of transforming a $(k+1)$-st order difference equation to a system of first order equations is simple and well-known (it is described in Section 3), the determination of the asymptotic properties of the solutions of the resulting system using either Liapunov approach or retract-type method is not trivial. These analogies of classical approaches, known from the qualitative theory of differential equations, were developed for difference systems in [1] (where an approach based on Liapunov method was formulated) and in [2–5] (where retract-type analysis was modified for discrete equations). It occurs that for the mentioned analysis of asymptotic problems of system (1.2), neither the ideas of Liapunov, nor the retract-type technique can be applied directly. However, in spite of the fact that each of the two mentioned methods fails when used independently, it appears that the combination of both these techniques works for this type of systems. Therefore, in Section 2 we prove the relevant result suitable for the asymptotic analysis of systems arising by transformation of (1.2) to a system of first-order differential equations (Theorem 2.1), where the assumptions put to the right-hand side of the system are of both types: those caused by the application of the Liapunov approach and those which are typical for the retract-type technique. Such an idea was applied in a particular case of investigation of asymptotic properties of solutions of the discrete analogue of the Emden-Fowler equation in [6, 7]. The approach is demonstrated in Section 3 where, moreover, its usefulness is illustrated on the problem of detecting the existence of positive solutions of linear equations with a single delay (in Section 3.4) and asymptotic estimation of solutions (in Section 3.3).

Advantages of our approach can be summarized as follows. We give a general method of analysis which is different from the well-known comparison method (see, e.g., [8, 9]). Comparing our approach with the scheme of investigation in [10, 11] which is based on a result from [12], we can see that the presented method is more general because it unifies

to Contents, page 1

the investigation of systems of discrete equations and delayed discrete equations thanks to the Liapunov-retract-type technique.

For related results concerning positive solutions and the asymptotics of solutions of discrete equations, the reader is referred also to [13–25].

2. The Result for Systems of First-Order Equations

Consider the system of m difference equations

$$\Delta u(n) = F(n, u(n)), \tag{2.1}$$

where $n \in \mathbb{Z}_a^\infty$, $u = (u_1, \dots, u_m)$, and $F : \mathbb{Z}_a^\infty \times \mathbb{R}^m \to \mathbb{R}^m$, $F = (F_1, \dots, F_m)$. The solution of system (2.1) is defined as a vector function $u : \mathbb{Z}_a^\infty \to \mathbb{R}^m$ such that for every $n \in \mathbb{Z}_a^\infty$, (2.1) is fulfilled. Again, if we prescribe initial conditions

$$u_i(a) = u_i^a \in \mathbb{R}, \quad i = 1, \dots, m \tag{2.2}$$

the initial problem (2.1), (2.2) has a unique solution. Let us define a set $\Omega \subset \mathbb{Z}_a^\infty \times \mathbb{R}^m$ as

$$\Omega := \bigcup_{n=a}^{\infty} \Omega(n), \tag{2.3}$$

where

$$\Omega(n) := \{(n, u) : n \in \mathbb{Z}_a^\infty,\ u_i \in \mathbb{R},\ b_i(n) < u_i < c_i(n),\ i = 1, \dots, m\} \tag{2.4}$$

with $b_i, c_i : \mathbb{Z}_a^\infty \to \mathbb{R}$, $i = 1, \dots, m$, being auxiliary functions such that $b_i(n) < c_i(n)$ for each $n \in \mathbb{Z}_a^\infty$. Such set Ω is called a *polyfacial set*.

Our aim (in this part) is to solve, in correspondence with formulated Problem 1, the following similar problem for systems of difference equations.

Problem 2. Derive sufficient conditions with respect to the right-hand sides of system (2.1) which guarantee the existence of at least one solution $u(n) = (u_1^*(n), \dots, u_m^*(n))$, $n \in \mathbb{Z}_a^\infty$, satisfying

$$\left(n, u_1^*(n), \dots, u_m^*(n)\right) \in \Omega(n) \tag{2.5}$$

for every $n \in \mathbb{Z}_a^\infty$.

As we mentioned above, in [1] the above described problem is solved via a Liapunov-type technique. Here we will combine this technique with the retract-type technique which was used in [2–5] so as the result can be applied easily to the system arising after transformation of (1.2). This brings a significant increase in the range of systems we are able to investigate. Before we start, we recall some basic notions that will be used.

to Contents, page 1

2.1. Consequent Point

Define the mapping $\mathcal{C} : \mathbb{Z}_a^\infty \times \mathbb{R}^m \to \mathbb{Z}_a^\infty \times \mathbb{R}^m$ as

$$\mathcal{C} : (n, u) \longmapsto (n + 1, u + F(n, u)). \tag{2.6}$$

For any point $M = (n, u) \in \mathbb{Z}_a^\infty \times \mathbb{R}^m$, the point $\mathcal{C}(M)$ is called the *first consequent point* of the point M. The geometrical meaning is that if a point M lies on the graph of some solution of system (2.1), then its first consequent point $\mathcal{C}(M)$ is the next point on this graph.

2.2. Liapunov-Type Polyfacial Set

We say that a polyfacial set Ω is *Liapunov-type* with respect to discrete system (2.1) if

$$b_i(n + 1) < u_i + F_i(n, u) < c_i(n + 1) \tag{2.7}$$

for every $i = 1, \ldots, m$ and every $(n, u) \in \Omega$. The geometrical meaning of this property is this: if a point $M = (n, u)$ lies inside the set $\Omega(n)$, then its first consequent point $\mathcal{C}(M)$ stays inside $\Omega(n + 1)$.

In this contribution we will deal with sets that need not be of Liapunov-type, but they will have, in a certain sense, a similar property. We say that a polyfacial set Ω is *Liapunov-type with respect to the jth variable* ($j \in \{1, \ldots, m\}$) and to discrete system (2.1) if

$$(n, u) \in \Omega \Longrightarrow b_j(n + 1) < u_j + F_j(n, u) < c_j(n + 1). \tag{2.8}$$

The geometrical meaning is that if $M = (n, u) \in \Omega(n)$, then the u_j-coordinate of its first consequent point stays between $b_j(n + 1)$ and $c_j(n + 1)$, meanwhile the other coordinates of $\mathcal{C}(M)$ may be arbitrary.

2.3. Points of Strict Egress and Their Geometrical Sense

An important role in the application of the retract-type technique is played by the so called strict egress points. Before we define these points, let us describe the boundaries of the sets $\Omega(n)$, $n \in \mathbb{Z}_a^\infty$, in detail. As one can easily see,

$$\bigcup_{n \in \mathbb{Z}_a^\infty} \partial\Omega(n) = \left(\bigcup_{j=1}^m \Omega_B^j \right) \cup \left(\bigcup_{j=1}^m \Omega_C^j \right) \tag{2.9}$$

with

$$\begin{aligned} \Omega_B^j &:= \{(n, u) : n \in \mathbb{Z}_a^\infty,\ u_j = b_j(n),\ b_i(n) \le u_i \le c_i(n),\ i = 1, \ldots, m,\ i \neq j\}, \\ \Omega_C^j &:= \{(n, u) : n \in \mathbb{Z}_a^\infty,\ u_j = c_j(n),\ b_i(n) \le u_i \le c_i(n),\ i = 1, \ldots, m,\ i \neq j\}. \end{aligned} \tag{2.10}$$

to Contents, page 1

In accordance with [3, Lemmas 1 and 2], a point $(n,u) \in \partial\Omega(n)$ is a *point of the type of strict egress* for the polyfacial set Ω with respect to discrete system (2.1) if and only if for some $j \in \{1,\ldots,m\}$

$$u_j = b_j(n), \quad F_j(n,u) < b_j(n+1) - b_j(n), \tag{2.11}$$

or

$$u_j = c_j(n), \quad F_j(n,u) > c_j(n+1) - c_j(n). \tag{2.12}$$

Geometrically these inequalities mean the following: if a point $M = (n,u) \in \partial\Omega(n)$ is a point of the type of strict egress, then the first consequent point $\mathcal{C}(M) \notin \overline{\Omega(n+1)}$.

2.4. Retract and Retraction

If $A \subset B$ are any two sets in a topological space and $\pi : B \to A$ is a continuous mapping from B onto A such that $\pi(p) = p$ for every $p \in A$, then π is said to be a *retraction* of B onto A. If there exists a retraction of B onto A, then A is called a *retract* of B.

2.5. The Existence Theorem for the System of First-Order Equations (Solution of Problem 2)

The following result, solving Problem 2, gives sufficient conditions with respect to the right-hand sides of (2.1) which guarantee the existence of at least one solution satisfying (2.5) for every $n \in \mathbb{Z}_a^\infty$.

Theorem 2.1. *Let $b_i(n), c_i(n), b_i(n) < c_i(n)$, $i = 1,\ldots,m$, be real functions defined on $\mathbb{Z}_a^\infty$ and let $F_i : \mathbb{Z}_a^\infty \times \mathbb{R}^m \to \mathbb{R}$, $i = 1,\ldots,m$, be continuous functions. Suppose that for one fixed $j \in \{1,\ldots,m\}$ all the points of the sets Ω_B^j, Ω_C^j are points of strict egress, that is, if $(n,u) \in \Omega_B^j$, then*

$$F_j(n,u) < b_j(n+1) - b_j(n), \tag{2.13}$$

and if $(n,u) \in \Omega_C^j$, then

$$F_j(n,u) > c_j(n+1) - c_j(n). \tag{2.14}$$

Further suppose that the set Ω is of Liapunov-type with respect to the ith variable for every $i \in \{1,\ldots,m\}$, $i \neq j$, that is, that for every $(n,u) \in \Omega$

$$b_i(n+1) < u_i + F_i(n,u) < c_i(n+1). \tag{2.15}$$

Then there exists a solution $u = (u_1^*(n),\ldots,u_m^*(n))$ of system (2.1) satisfying the inequalities

$$b_i(n) < u_i^*(n) < c_i(n), \quad i = 1,\ldots,m, \tag{2.16}$$

for every $n \in \mathbb{Z}_a^\infty$.

to Contents, page 1

Proof. The proof will be by contradiction. We will suppose that there exists no solution satisfying inequalities (2.16) for every $n \in \mathbb{Z}_a^{\infty}$. Under this supposition we prove that there exists a continuous mapping (a retraction) of a closed interval onto both its endpoints which is, by the intermediate value theorem of calculus, impossible.

Without the loss of generality we may suppose that the index j in Theorem 2.1 is equal to 1, that is, all the points of the sets Ω_B^1 and Ω_C^1 are strict egress points. Each solution of system (2.1) is uniquely determined by the chosen initial condition

$$u(a) = (u_1(a), \ldots, u_m(a)) = \left(u_1^a, \ldots, u_m^a\right) = u^a. \tag{2.17}$$

For the following considerations, let u_i^a with $u_i^a \in (b_i(a), c_i(a))$, $i = 2, \ldots, m$, be chosen arbitrarily but fixed. Now the solution of (2.1) is given just by the choice of u_1^a, we can write

$$u = u\left(n, u_1^a\right) = \left(u_1\left(n, u_1^a\right), \ldots, u_m\left(n, u_1^a\right)\right). \tag{2.18}$$

Define the closed interval $I := [b_1(a), c_1(a)]$. Hereafter we show that, under the supposition that there exists no solution satisfying inequalities (2.16), there exists a retraction $\mathcal{R}$ (which will be a composition of two auxiliary mappings $\mathcal{R}_1$ and $\mathcal{R}_2$ defined below) of the set $B := I$ onto the set $A := \partial I = \{b_1(a), c_1(a)\}$. This contradiction will prove our result. To arrive at such a contradiction, we divide the remaining part of the proof into several steps.

Construction of the Leaving Value n^*

Let a point $\tilde{u}_1 \in I$ be fixed. The initial condition $u_1(a) = \tilde{u}_1$ defines a solution $u = u(n, \tilde{u}_1) = (u_1(n, \tilde{u}_1), \ldots, u_m(n, \tilde{u}_1))$. According to our supposition, this solution does not satisfy inequalities (2.16) for every $n \in \mathbb{Z}_a^{\infty}$. We will study the moment the solution leaves the domain Ω for the first time. The first value of n for which inequalities (2.16) are not valid will be denoted as s.

(I) First consider the case $\tilde{u}_1 \in \operatorname{int} I$. Then there exists a value $s > 1$ in $\mathbb{Z}_{a+1}^{\infty}$ such that

$$(s, u(s, \tilde{u}_1)) \notin \Omega(s) \tag{2.19}$$

while

$$(r, u(r, \tilde{u}_1)) \in \Omega(r) \quad \text{for } a \le r \le s-1. \tag{2.20}$$

As the set Ω is of the Liapunov-type with respect to all variables except the first one and $(s-1, u(s-1, \tilde{u}_1)) \in \Omega(s-1)$, then

$$b_i(s) < u_i(s, \tilde{u}_1) < c_i(s), \quad i = 2, \ldots, m. \tag{2.21}$$

Because $j = 1$ was assumed, and Ω is of Liapunov-type for each variable u_i, $i \neq j$, then the validity of inequalities (2.16) has to be violated in the u_1-coordinate. The geometrical meaning was explained in Section 2.2.

Now, two cases are possible: either $(s, u(s, \tilde{u}_1)) \notin \overline{\Omega(s)}$ or $(s, u(s, \tilde{u}_1)) \in \partial\Omega(s)$. In the first case $u_1(s, \tilde{u}_1) < b_1(s)$ or $u_1(s, \tilde{u}_1) > c_1(s)$. In the second case $u_1(s, \tilde{u}_1) = b_1(s)$ or

to Contents, page 1

$u_1(s,\tilde{u}_1) = c_1(s)$ and, due to (2.13) and (2.14), $u_1(s+1,\tilde{u}_1) < b_1(s+1)$ or $u_1(s+1,\tilde{u}_1) > c_1(s+1)$, respectively.

(II) If $\tilde{u}_1 \in \partial I$, then $(a, u(a,\tilde{u}_1)) \notin \Omega(a)$. Thus, for this case, we could put $s = a$. Further, because of the strict egress property of Ω^1_B and Ω^1_C, either $u_1(a+1,\tilde{u}_1) < b_1(a+1)$ (if $\tilde{u}_1 = b_1(a)$) or $u_1(a+1,\tilde{u}_1) > c_1(a+1)$ (if $\tilde{u}_1 = c_1(a)$) and thus $(a+1, u(a+1,\tilde{u}_1)) \notin \overline{\Omega(a+1)}$.

Unfortunately, for the next consideration the value s (the first value of the independent variable for which the graph of the solution is out of Ω) would be of little use. What we will need is the last value for which the graph of the solution stays in $\overline{\Omega}$. We will denote this value as n^* and will call it the *leaving value*. We can define n^* as

$$\begin{aligned} n^* &= s-1 \quad \text{if } (s, u(s,\tilde{u}_1)) \notin \overline{\Omega(s)}, \\ n^* &= s \quad \text{if } (s, u(s,\tilde{u}_1)) \in \partial\Omega(s). \end{aligned} \tag{2.22}$$

As the value of n^* depends on the chosen initial point $\tilde{u}_1$, we could write $n^* = n^*(\tilde{u}_1)$ but we will mostly omit the argument $\tilde{u}_1$, unless it is necessary. From the above considerations it follows that

$$\begin{gathered} b_1(n^*) \le u_1(n^*,\tilde{u}_1) \le c_1(n^*), \\ u_1(n^*+1,\tilde{u}_1) < b_1(n^*+1) \quad \text{or} \quad u_1(n^*+1,\tilde{u}_1) > c_1(n^*+1). \end{gathered} \tag{2.23}$$

Auxiliary Mapping $\mathcal{R}_1$

Now we construct the auxiliary mapping $\mathcal{R}_1 : I \to \mathbb{R} \times \mathbb{R}$. First extend the discrete functions b_1, c_1 onto the whole interval $[a,\infty)$:

$$\begin{aligned} b_1(t) &:= b_1(\lfloor t \rfloor) + (b_1(\lfloor t \rfloor + 1) - b_1(\lfloor t \rfloor))(t - \lfloor t \rfloor), \\ c_1(t) &:= c_1(\lfloor t \rfloor) + (c_1(\lfloor t \rfloor + 1) - c_1(\lfloor t \rfloor))(t - \lfloor t \rfloor), \end{aligned} \tag{2.24}$$

$\lfloor t \rfloor$ being the integer part of t (the floor function). Note that b_1, c_1 are now piecewise linear continuous functions of a real variable t such that $b_1(t) < c_1(t)$ for every t and that the original values of $b_1(n), c_1(n)$ for $n \in \mathbb{Z}^\infty_a$ are preserved. This means that the graphs of these functions connect the points $(n, b_1(n))$ or $(n, c_1(n))$ for $n \in \mathbb{Z}^\infty_a$, respectively. Denote V the set

$$V := \{(t, u_1) : t \in [a,\infty),\ b_1(t) \le u_1 \le c_1(t)\}. \tag{2.25}$$

The boundary of V consists of three mutually disjoint parts V_a, V_b, and V_c:

$$\partial V = V_a \cup V_b \cup V_c, \tag{2.26}$$

where

$$\begin{aligned} V_a &:= \{(a, u_1) : b_1(a) < u_1 < c_1(a)\}, \\ V_b &:= \{(t, u_1) : t \in [a,\infty), u_1 = b_1(t)\}, \\ V_c &:= \{(t, u_1) : t \in [a,\infty), u_1 = c_1(t)\}. \end{aligned} \tag{2.27}$$

to Contents, page 1

Define the mapping $\mathcal{R}_1 : I \to V_b \cup V_c$ as follows: let $\mathcal{R}_1(\tilde{u}_1)$ be the point of intersection of the line segment defined by its end points $(n^*, u_1(n^*, \tilde{u}_1))$, $(n^*+1, u_1(n^*+1, \tilde{u}_1))$ with $V_b \cup V_c$ (see Figure 1). The mapping $\mathcal{R}_1$ is obviously well defined on I and $\mathcal{R}_1(b_1(a)) = (a, b_1(a))$, $\mathcal{R}_1(c_1(a)) = (a, c_1(a))$.

Prove that the mapping $\mathcal{R}_1$ is continuous. The point $\mathcal{R}_1(\tilde{u}_1) = (t(\tilde{u}_1), u_1(\tilde{u}_1))$ lies either on V_b or on V_c. Without the loss of generality, consider the second case (the first one is analogical). The relevant boundary line segment for $t \in [n^*, n^*+1]$, which is a part of V_c, is described by (see (2.24))

$$u_1 = c(n^*) + (c(n^*+1) - c(n^*))(t - n^*), \tag{2.28}$$

and the line segment joining the points $(n^*, u_1(n^*, \tilde{u}_1))$, $(n^*+1, u_1(n^*+1, \tilde{u}_1))$ by the equation

$$u_1 = u_1(n^*, \tilde{u}_1) + (u_1(n^*+1, \tilde{u}_1) - u_1(n^*, \tilde{u}_1))(t - n^*), \quad t \in [n^*, n^*+1]. \tag{2.29}$$

The coordinates of the point $\mathcal{R}_1(\tilde{u}_1) = (t(\tilde{u}_1), u_1(\tilde{u}_1))$, which is the intersection of both these line segments, can be obtained as the solution of the system consisting of (2.28) and (2.29). Solving this system with respect to t and u_1, we get

$$t(\tilde{u}_1) = n^* + \frac{u_1(n^*, \tilde{u}_1) - c_1(n^*)}{c_1(n^*+1) - u_1(n^*+1, \tilde{u}_1) + u_1(n^*, \tilde{u}_1) - c_1(n^*)}, \tag{2.30}$$

$$u_1(\tilde{u}_1) = c_1(n^*) + \frac{(u_1(n^*, \tilde{u}_1) - c_1(n^*))(c_1(n^*+1) - c_1(n^*))}{c_1(n^*+1) - u_1(n^*+1, \tilde{u}_1) + u_1(n^*, \tilde{u}_1) - c_1(n^*)}. \tag{2.31}$$

Let $\{v_k\}_{k=1}^{\infty}$ be any sequence with $v_k \in I$ such that $v_k \to \tilde{u}_1$. We will show that $\mathcal{R}_1(v_k) \to \mathcal{R}_1(\tilde{u}_1)$. Because of the continuity of the functions F_i, $i = 1, \ldots, m$,

$$u_1(n, v_k) \to u_1(n, \tilde{u}_1) \quad \text{for every fixed } n \in \mathbb{Z}_a^{\infty}. \tag{2.32}$$

We have to consider two cases:

(I) $(n^*, u(n^*, \tilde{u}_1)) \in \Omega(n^*)$, that is, $b_1(n^*) < u_1(n^*, \tilde{u}_1) < c_1(n^*)$,

(II) $(n^*, u(n^*, \tilde{u}_1)) \in \partial\Omega(n^*)$, that is, $u_1(n^*, \tilde{u}_1) = c_1(n^*)$.

Recall that (due to our agreement) in both cases $u_1(n^*+1, \tilde{u}_1) > c_1(n^*+1)$.

(I) In this case also $u_1(n^*, v_k) < c_1(n^*)$ and $u_1(n^*+1, v_k) > c_1(n^*+1)$ for k sufficiently large. That means that the leaving value $n^*(v_k)$ is the same as n^* given by $\tilde{u}_1$ and thus the point $\mathcal{R}_1(v_k) = (t(v_k), u_1(v_k))$ is given by

$$t(v_k) = n^* + \frac{u_1(n^*, v_k) - c_1(n^*)}{c_1(n^*+1) - u_1(n^*+1, v_k) + u_1(n^*, v_k) - c_1(n^*)}, \tag{2.33}$$

$$u_1(v_k) = c_1(n^*) + \frac{(u_1(n^*, v_k) - c_1(n^*))(c_1(n^*+1) - c_1(n^*))}{c_1(n^*+1) - u_1(n^*+1, v_k) + u_1(n^*, v_k) - c_1(n^*)}. \tag{2.34}$$

to Contents, page 1

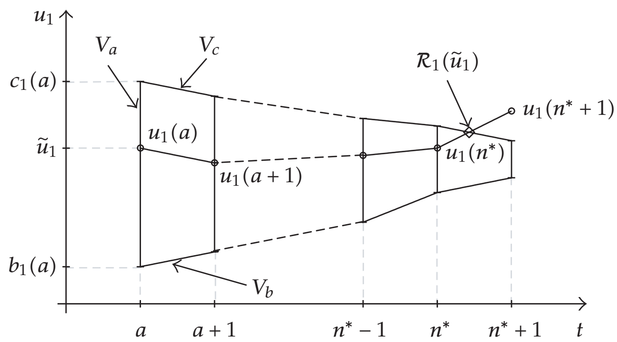


Figure 1: Construction of the mapping $\mathcal{R}_1$.

The desired convergence $\mathcal{R}_1(v_k) \to \mathcal{R}_1(\tilde{u}_1)$ is implied by equations (2.30) to (2.34).

(II) Suppose $n^* = a$. Then $\tilde{u}_1 = c_1(a)$, $v_k = u_1(a, v_k) < c_1(a)$ for all k and as $k \to \infty$, $u_1(a+1, v_k) > c_1(a+1)$. A minor edit of the text in the case (I) proof provides the continuity proof.

Suppose $n^* > a$. In this case there can be $u_1(n^*, v_k) \leq c_1(n^*)$ for some members of the sequence $\{v_k\}$ and $u_1(n^*, v_k) > c_1(n^*)$ for the others. Without the loss of generality, we can suppose that $\{v_k\}$ splits into two infinite subsequences $\{v_{q_k}\}$ and $\{v_{r_k}\}$ such that

$$u_1(n^*, v_{q_k}) \leq c_1(n^*), \qquad u_1(n^*+1, v_{q_k}) > c_1(n^*+1)$$
$$u_1(n^*, v_{r_k}) > c_1(n^*). \tag{2.35}$$

For the subsequence $\{v_{q_k}\}$, the text of the proof of (I) can be subjected to a minor edit to provide the proof of continuity. As for the subsequence $\{v_{r_k}\}$, the leaving value $n^*(v_{r_k})$ is different from n^* given by $\tilde{u}_1$ because $(n^*, u_1(n^*, v_{r_k}))$ is already out of $\overline{\Omega}$. For k sufficiently large,

$$n^*(v_{r_k}) = n^* - 1 \tag{2.36}$$

because $u_1(n^*-1, \tilde{u}_1) < c_1(n^*-1)$ and thus, as $k \to \infty$, $u_1(n^*-1, v_{r_k}) < c_1(n^*-1)$. Hence, the value of the mapping $\mathcal{R}_1$ for v_{r_k} is (in (2.33), (2.34) we replace n^* by n^*-1)

$$t(v_{r_k}) = n^* - 1 + \frac{u_1(n^*-1, v_{r_k}) - c_1(n^*-1)}{c_1(n^*) - u_1(n^*, v_{r_k}) + u_1(n^*-1, v_{r_k}) - c_1(n^*-1)},$$
$$u_1(v_{r_k}) = c(n^*-1) + \frac{(u_1(n^*-1, v_{r_k}) - c_1(n^*-1))(c_1(n^*) - c_1(n^*-1))}{c_1(n^*) - u_1(n^*, v_{r_k}) + u_1(n^*-1, v_{r_k}) - c_1(n^*-1)}. \tag{2.37}$$

to Contents, page 1

Due to (2.32), $u_1(n^*, v_{r_k}) \to u_1(n^*, \tilde{u}_1) = c_1(n^*)$ and thus

$$t(v_{r_k}) \longrightarrow n^* - 1 + \frac{u_1(n^*-1, v_{r_k}) - c_1(n^*-1)}{u_1(n^*-1, v_{r_k}) - c_1(n^*-1)} = n^*,$$
$$u_1(v_{r_k}) \longrightarrow c(n^*-1) + \frac{(u_1(n^*-1, v_{r_k}) - c_1(n^*-1))(c_1(n^*) - c_1(n^*-1))}{u_1(n^*-1, v_{r_k}) - c_1(n^*-1)} = c_1(n^*), \tag{2.38}$$
$$\mathcal{R}_1(v_{r_k}) = (t(v_{r_k}), u_1(v_{r_k})) \longrightarrow (n^*, c_1(n^*)) = \mathcal{R}_1(\tilde{u}_1).$$

We have shown that $\mathcal{R}_1(v_{q_k}) \to \mathcal{R}_1(\tilde{u}_1)$ and $\mathcal{R}_1(v_{r_k}) \to \mathcal{R}_1(\tilde{u}_1)$ and thus $\mathcal{R}_1(v_k) \to \mathcal{R}_1(\tilde{u}_1)$.

Auxiliary Mapping $\mathcal{R}_2$

Define $\mathcal{R}_2 : V_b \cup V_c \to \{b_1(a), c_1(a)\}$ as

$$\mathcal{R}_2(P) = \begin{cases} b_1(a) & \text{if } P \in V_b, \\ c_1(a) & \text{if } P \in V_c. \end{cases} \tag{2.39}$$

The mapping $\mathcal{R}_2$ is obviously continuous.

Resulting Mapping $\mathcal{R}$ and Its Properties

Define $\mathcal{R} := \mathcal{R}_2 \circ \mathcal{R}_1$. Due to construction we have

$$\mathcal{R}(b_1(a)) = b_1(a), \ \mathcal{R}(c_1(a)) = c_1(a), \tag{2.40}$$

and $\mathcal{R}(I) = \partial I$. The mapping $\mathcal{R}$ is continuous because of the continuity of the two mappings $\mathcal{R}_1$ and $\mathcal{R}_2$. Hence, it is the sought retraction of I onto ∂I. But such a retraction cannot exist and thus we get a contradiction and the proof is complete. □

3. Application of Theorem 2.1 to the Delayed Discrete Equation

Now, let us return to the original delayed discrete equation (1.2), that is,

$$\Delta v(n) = f(n, v(n), v(n-1), \ldots, v(n-k)). \tag{3.1}$$

As it was said in Section 1, this equation will be transformed to a system of $k+1$ first-order discrete equations. Then we will apply Theorem 2.1 to this system and prove that under certain conditions there exists a solution of delayed equation (1.2) that stays in the prescribed domain. In the end, we will study a special case of (1.2).

3.1. Transformation of (1.2) to the System of First-Order Equations

We will proceed in accordance with the well-known scheme similarly as when constructing the system of first-order differential equations from a differential equation of a higher order. Put

$$\begin{aligned} u_1(n) &:= v(n), \\ u_2(n) &:= v(n-1), \\ &\cdots \\ u_{k+1}(n) &:= v(n-k), \end{aligned} \tag{3.2}$$

where $u_1, u_2, \ldots, u_{k+1}$ are new unknown functions. From (1.2) we get $\Delta u_1(n) = f(n, u_1(n), u_2(n), \ldots, u_{k+1}(n))$. Obviously $u_2(n+1) = u_1(n), \ldots, u_{k+1}(n+1) = u_k(n)$. Rewriting these equalities in terms of differences, we have $\Delta u_2(n) = u_1(n) - u_2(n), \ldots, \Delta u_{k+1}(n) = u_k(n) - u_{k+1}(n)$. Altogether, we get the system

$$\begin{aligned} \Delta u_1(n) &= f(n, u_1(n), \ldots, u_{k+1}(n)), \\ \Delta u_2(n) &= u_1(n) - u_2(n), \\ &\cdots \\ \Delta u_{k+1}(n) &= u_k(n) - u_{k+1}(n) \end{aligned} \tag{3.3}$$

which is equivalent to (1.2).

3.2. The Existence Theorem for the Delayed Equation (1.2) (Solution of Problem 1)

The following theorem is a consequence of Theorem 2.1. In fact, this theorem has been already proved in [12]. There, the proof is based upon a modification of the retract method for delayed equations. Our method (rearranging a delayed equation to a system of first-order equations) is, by its principle, more general than that used in [12].

Theorem 3.1. *Let $b(n)$, $c(n)$, $b(n) < c(n)$, be real functions defined on $\mathbb{Z}_{a-k}^{\infty}$. Further, let $f : \mathbb{Z}_a^{\infty} \times \mathbb{R}^{k+1} \to \mathbb{R}$ be a continuous function and let the inequalities*

$$b(n) + f(n, b(n), v_2, \ldots, v_{k+1}) < b(n+1), \tag{3.4}$$

$$c(n) + f(n, c(n), v_2, \ldots, v_{k+1}) > c(n+1) \tag{3.5}$$

hold for every $n \in \mathbb{Z}_a^{\infty}$ and every $v_2, \ldots, v_{k+1}$ such that

$$b(n-i+1) < v_i < c(n-i+1), \quad i = 2, \ldots, k+1. \tag{3.6}$$

to Contents, page 1

Then there exists a solution $v = v^*(n)$ of (1.2) *satisfying the inequalities*

$$b(n) < v^*(n) < c(n) \tag{3.7}$$

for every $n \in \mathbb{Z}_{a-k}^{\infty}$.

Proof. We have shown that (1.2) is equivalent to system (3.3) which can be seen as a special case of system (2.1) with $m = k+1$ and $F = (F_1, \dots F_{k+1})$ where

$$\begin{gathered} F_1(n, u_1, \dots, u_{k+1}) := f(n, u_1, \dots, u_{k+1}), \\ F_2(n, u_1, \dots, u_{k+1}) := u_1 - u_2, \\ \cdots \\ F_k(n, u_1, \dots, u_{k+1}) := u_{k-1} - u_k, \\ F_{k+1}(n, u_1, \dots, u_{k+1}) := u_k - u_{k+1}. \end{gathered} \tag{3.8}$$

Define the polyfacial set Ω as

$$\Omega := \{(n, u) : n \in \mathbb{Z}_a^{\infty},\ b_i(n) < u_i < c_i(n),\ i = 1, \dots, k+1\} \tag{3.9}$$

with

$$b_i(n) := b(n - i + 1), \qquad c_i(n) := c(n - i + 1), \quad i = 1, \dots, k+1. \tag{3.10}$$

We will show that for system (3.3) and the set Ω, all the assumptions of Theorem 2.1 are satisfied.

As the function f is supposed to be continuous, the mapping F is continuous, too. Put the index j from Theorem 2.1, characterizing the points of egress, equal to 1. We will verify that the set Ω is of Liapunov-type with respect to the ith variable for any $i = 2, \dots, k+1$, that is, (see (2.8)) that for every $(n, u) \in \Omega$

$$b_i(n+1) < u_i + F_i(n, u) < c_i(n+1) \quad \text{for } i = 2, \dots, k+1. \tag{3.11}$$

First, we compute

$$u_i + F_i(n, u) = u_i + u_{i-1} - u_i = u_{i-1} \quad \text{for } i = 2, \dots, k+1. \tag{3.12}$$

Thus we have to show that for $i = 2, \dots, k+1$

$$b_i(n+1) < u_{i-1} < c_i(n+1). \tag{3.13}$$

Because $(n, u) \in \Omega$, then $b_p(n) < u_p < c_p(n)$ for any $p \in \{1, \dots, k+1\}$, and therefore

$$b_{i-1}(n) < u_{i-1} < c_{i-1}(n) \quad \text{for } i = 2, \dots, k+1. \tag{3.14}$$

to Contents, page 1

But, by (3.10), we have

$$b_{i-1}(n) = b(n-i+1+1) = b(n-i+2), \tag{3.15}$$

meanwhile

$$b_i(n+1) = b(n+1-i+1) = b(n-i+2), \tag{3.16}$$

and thus $b_{i-1}(n) = b_i(n+1)$. Analogously we get that $c_{i-1}(n) = c_i(n+1)$. Thus inequalities (3.11) are fulfilled.

Further we will show that all the boundary points $M \in \Omega_B^1 \cup \Omega_C^1$ are points of strict egress for the set Ω with respect to system (3.3). According to (2.11), we have to show that if $u_1 = b_1(n)$ and $b_i(n) < u_i < c_i(n)$ for $i = 2, \dots, k+1$, then

$$b_1(n) + F_1(n,u) < b_1(n+1), \tag{3.17}$$

that is,

$$b_1(n) + f(n, b_1(n), u_2, \dots, u_{k+1}) < b_1(n+1). \tag{3.18}$$

Notice that the condition $b_i(n) < u_i < c_i(n)$ for $i = 2, \dots, k+1$ is equivalent with condition $b(n-i+1) < u_i < c(n-i+1)$ (see (3.10)). Looking at the supposed inequality (3.4) and realizing that $b_1(n) = b(n)$ and $b_1(n+1) = b(n+1)$, we can see that inequality (3.18) is fulfilled.

Analogously, according to (2.12), we have to prove that for $u_1 = c_1(n)$ and $b_i(n) < u_i < c_i(n)$ for $i = 2, \dots, k+1$ the inequality

$$c_1(n) + F_1(n,u) > c_1(n+1), \tag{3.19}$$

that is,

$$c_1(n) + f(n, c_1(n), u_2, \dots, u_{k+1}) > c_1(n+1) \tag{3.20}$$

holds.

Again, considering (3.5) and the fact that $c_1(n) = c(n)$ and $c_1(n+1) = c(n+1)$, we can see that this inequality really holds.

Thus, by the assertion of Theorem 2.1, there exists a solution $u = u^*(n)$ of system (3.3) such that for every $n \in \mathbb{Z}_a^\infty$

$$b_i(n) < u_i^*(n) < c_i(n) \quad \text{for } i = 1, \dots, k+1. \tag{3.21}$$

In our case, $v = v^*(n) = u_1^*(n)$ is the solution of the original equation (1.2). Further, $b_1(n) = b(n)$ and $c_1(n) = c(n)$, and thus the existence of a solution of the delayed equation (1.2) such that inequalities (3.7) are satisfied is guaranteed. □

to Contents, page 1

3.3. Asymptotic Solution Estimates for Delayed Difference Equations

Let us suppose that two functions $u, w : \mathbb{Z}_{a-k}^{\infty} \to \mathbb{R}$ are given such that

$$u(n) < w(n), \quad n \in \mathbb{Z}_{a-k}^{\infty}, \tag{3.22}$$

$$\Delta u(n) \geq f(n, u(n), u(n-1), \ldots, u(n-k)), \quad n \in \mathbb{Z}_{a}^{\infty}, \tag{3.23}$$

$$\Delta w(n) \leq f(n, w(n), w(n-1), \ldots, w(n-k)), \quad n \in \mathbb{Z}_{a}^{\infty}. \tag{3.24}$$

Consider the problem of whether there exists a solution $v = v^*(n)$, $n \in \mathbb{Z}_{a-k}^{\infty}$ of (1.2) such that

$$u(n) < v^*(n) < w(n), \quad n \in \mathbb{Z}_{a-k}^{\infty}. \tag{3.25}$$

The following corollary of Theorem 3.1 presents sufficient conditions for the existence of a solution of this problem.

Corollary 3.2. *Let functions $u, w : \mathbb{Z}_{a-k}^{\infty} \to \mathbb{R}$ satisfy inequalities* (3.22)–(3.24). *Let $f : \mathbb{Z}_{a}^{\infty} \times \mathbb{R}^{k+1} \to \mathbb{R}$ be a continuous function such that*

$$f(n, u(n), y_2, \ldots, y_{k+1}) > f(n, u(n), z_2, \ldots, z_{k+1}), \tag{3.26}$$

$$f(n, w(n), y_2, \ldots, y_{k+1}) > f(n, w(n), z_2, \ldots, z_{k+1}) \tag{3.27}$$

for every $n \in \mathbb{Z}_{a}^{\infty}$ and every $y_2, \ldots, y_{k+1}, z_2, \ldots, z_{k+1} \in \mathbb{R}$ such that

$$y_i < z_i, \quad i = 2, \ldots, k+1. \tag{3.28}$$

Then there exists a solution $v = v^*(n)$ of (1.2) *satisfying inequalities* (3.25) *for every $n \in \mathbb{Z}_{a-k}^{\infty}$.*

Proof. This assertion is an easy consequence of Theorem 3.1.

Put $b(n) := u(n)$, $c(n) := w(n)$. Considering inequalities (3.23) and (3.26), we can see that

$$\Delta u(n) > f(n, u(n), v_2, \ldots, v_{k+1}) \tag{3.29}$$

for every $n \in \mathbb{Z}_{a}^{\infty}$ and every $v_2, \ldots, v_{k+1}$ such that

$$b(n-i+1) < v_i < c(n-i+1), \quad i = 2, \ldots, k+1. \tag{3.30}$$

Similarly,

$$\Delta w(n) < f(n, w(n), v_2, \ldots, v_{k+1}) \tag{3.31}$$

for every $n \in \mathbb{Z}_{a}^{\infty}$ and every $b(n-i+1) < v_i < c(n-i+1)$, $i = 2, \ldots, k+1$.

to Contents, page 1

Obviously, inequalities (3.29) and (3.31) are equivalent with inequalities (3.4) and (3.5), respectively. Thus, all the assumptions of Theorem 3.1 are satisfied and there exists a solution $v = v^*(n)$ of (1.2) satisfying inequalities (3.25) for every $n \in \mathbb{Z}_{a-k}^{\infty}$. □

Example 3.3. Consider the equation

$$\Delta v(n) = v^2(n) - v(n-1) \tag{3.32}$$

for $n \in \mathbb{Z}_3^{\infty}$ which is a second-order delayed discrete equation with delay $k = 1$. We will show that there exists a solution $v = v^*(n)$ of (3.32) that satisfies the inequalities

$$1 < v^*(n) < n \tag{3.33}$$

for $n \in \mathbb{Z}_2^{\infty}$.

We will prove that for the functions

$$u(n) := 1, \quad w(n) := n, \quad f(n, v_1, v_2) := v_1^2 - v_2 \tag{3.34}$$

all the assumptions of Corollary 3.2 are satisfied. Inequality (3.22) is obviously fulfilled for $n \in \mathbb{Z}_2^{\infty}$. Inequality (3.23) can be also proved very easily:

$$\Delta u(n) = 0, \qquad f(n, u(n), u(n-1)) = 1^2 - 1 = 0, \tag{3.35}$$

and thus for every $n \in \mathbb{Z}_3^{\infty}$, $\Delta u(n) \geq f(n, u(n), u(n-1))$.

As for inequality (3.24), we get

$$\Delta w(n) = 1, \qquad f(n, w(n), w(n-1)) = n^2 - n + 1 \tag{3.36}$$

and thus $\Delta w(n) \leq f(n, w(n), w(n-1))$ for $n \in \mathbb{Z}_3^{\infty}$.

Finally, the functions

$$f(n, u(n), v_2) = 1 - v_2, \qquad f(n, w(n), v_2) = n^2 - v_2 \tag{3.37}$$

are decreasing with respect to v_2. Therefore, conditions (3.26) and (3.27) are satisfied, too. Hence, due to Corollary 3.2, there exists a solution of (3.32) satisfying (3.33).

3.4. Positive Solutions of a Linear Equation with a Single Delay

We will apply the result of Theorem 3.1 to the investigation of a simple linear difference equation of the $(k+1)$-st order with only one delay, namely, the equation

$$\Delta v(n) = -p(n)v(n-k), \tag{3.38}$$

where, again, $n \in \mathbb{Z}_a^{\infty}$ is the independent variable and $k \in \mathbb{N}$, $k \geq 1$, is the fixed delay. The function $p : \mathbb{Z}_a^{\infty} \to \mathbb{R}$ is assumed to be positive. Our goal is to give sharp sufficient conditions

to Contents, page 1

for the existence of positive solutions. The existence of such solutions is very often substantial for a concrete model considered. For example, in biology, when a model of population dynamics is described by an equation, the positivity of a solution may mean that the studied biological species can survive in the supposed environment.

For its simple form, (3.38) often serves for testing new results and is very frequently investigated. It was analyzed, for example, in papers [10, 11, 26]. A sharp result on existence of positive solutions given in [26] is proved by a comparison method [8, 9]. Here we will use Theorem 3.1 to generalize this result.

For the purposes of this section, define the expression $\ln_q t$, where $q \in \mathbb{N}$, as

$$\begin{aligned} \ln_q t &:= \ln(\ln_{q-1} t) \\ \ln_0 t &:= t. \end{aligned} \tag{3.39}$$

We will write only $\ln t$ instead of $\ln_1 t$. Further, for a fixed integer $\ell \geq 0$ define auxiliary functions

$$\begin{aligned} \mu_\ell(n) &:= \frac{1}{8n^2} + \frac{1}{8(n\ln n)^2} + \cdots + \frac{1}{8(n\ln n \cdots \ln_\ell n)^2}, \\ p_\ell(n) &:= \left(\frac{k}{k+1}\right)^k \cdot \left(\frac{1}{k+1} + k\mu_\ell(n)\right), \end{aligned} \tag{3.40}$$

$$\nu_\ell(n) := \left(\frac{k}{k+1}\right)^n \cdot \sqrt{n \ln n \ln_2 n \cdots \ln_\ell n}. \tag{3.41}$$

In [26], it was proved that if $p(n)$ in (3.38) is a positive function bounded by $p_\ell(n)$ for some $\ell \geq 0$, then there exists a positive solution of (3.38) bounded by the function $\nu_\ell(n)$ for n sufficiently large. Since $\lim_{n\to\infty} \nu_\ell(n) = 0$, such solution will vanish for $n \to \infty$. Here we show that (3.38) has a positive solution bounded by $\nu_\ell(n)$ even if the coefficient $p(n)$ satisfies a less restrictive inequality (see inequality (3.58) below). The proof of this statement will be based on the following four lemmas. The symbols "o" and "O" stand for the Landau order symbols and are used for $n \to \infty$.

Lemma 3.4. *The formula*

$$\ln(y - z) = \ln y - \sum_{i=1}^{\infty} \frac{z^i}{iy^i} \tag{3.42}$$

holds for any numbers $y, z \in \mathbb{R}$ such that $y > 0$ and $|z| < y$.

Proof. The assertion is a simple consequence of the well-known Maclaurin expansion

$$\ln(1 - x) = -\sum_{i=1}^{\infty} \frac{1}{i} x^i \quad \text{for } -1 \leq x < 1. \tag{3.43}$$

to Contents, page 1

As $\ln(y-z) - \ln y = \ln(1 - z/y)$, substituting $x = z/y$ we get

$$\ln(y-z) - \ln y = -\sum_{i=1}^{\infty} \frac{z^i}{iy^i} \quad \text{for} \ -y \le z < y \tag{3.44}$$

and adding $\ln y$ to both sides of this equality, we get (3.42). □

Lemma 3.5. *For fixed $r \in \mathbb{R} \setminus \{0\}$ and fixed $q \in \mathbb{N}$, the asymptotic representation*

$$\begin{aligned}\ln_q(n-r) = \ln_q n &- \frac{r}{n \ln n \cdots \ln_{q-1} n} - \frac{r^2}{2n^2 \ln n \cdots \ln_{q-1} n} \\ &- \frac{r^2}{2(n \ln n)^2 \ln_2 n \cdots \ln_{q-1} n} - \cdots - \frac{r^2}{2(n \ln n \cdots \ln_{q-1} n)^2} \\ &- \frac{r^3(1+o(1))}{3n^3 \ln n \cdots \ln_{q-1} n}\end{aligned} \tag{3.45}$$

holds for $n \to \infty$.

Proof. We will prove relation (3.45) by induction with respect to q. For $q = 1$, (3.45) reduces to

$$\ln(n-r) = \ln n - \frac{r}{n} - \frac{r^2}{2n^2} - \frac{r^3(1+o(1))}{3n^3} \tag{3.46}$$

which holds due to Lemma 3.4. Suppose that relation (3.45) holds for some q. We can write $\ln_q(n-r) = y - z$ with $y = \ln_q n$ and

$$\begin{aligned}z = &\frac{r}{n \ln n \cdots \ln_{q-1} n} + \frac{r^2}{2n^2 \ln n \cdots \ln_{q-1} n} + \frac{r^2}{2(n \ln n)^2 \ln_2 n \cdots \ln_{q-1} n} \\ &+ \cdots + \frac{r^2}{2(n \ln n \cdots \ln_{q-1} n)^2} + \frac{r^3(1+o(1))}{3n^3 \ln n \cdots \ln_{q-1} n}.\end{aligned} \tag{3.47}$$

Now we will show that (3.45) holds for $q+1$. Notice that in our case, the condition $|z| < y$ from Lemma 3.4 is fulfilled for n sufficiently large because $z \to 0$ for $n \to \infty$, meanwhile $y \to \infty$ for $n \to \infty$. Thus we are justified to use Lemma 3.4 and doing so, we get

to Contents, page 1

$$
\begin{aligned}
\ln_{q+1}(n-r) &= \ln\left(\ln_q(n-r)\right) \\
&= \ln(y-z) = \ln y - \frac{1}{y}\, z - \frac{1}{2y^2}\, z^2 - \cdots \\
&= \ln(\ln_q n) - \frac{1}{\ln_q n}\cdot\left(\frac{r}{n\ln n\cdots\ln_{q-1}n} + \frac{r^2}{2n^2\ln n\cdots\ln_{q-1}n} + \cdots \right. \\
&\qquad\qquad \left. + \frac{r^2}{2\left(n\ln n\cdots\ln_{q-1}n\right)^2} + \frac{r^3(1+o(1))}{3n^3\ln n\cdots\ln_{q-1}n}\right) \\
&\quad - \frac{1}{2\left(\ln_q n\right)^2}\cdot\left(\frac{r^2}{\left(n\ln n\cdots\ln_{q-1}n\right)^2} + O\left(\frac{1}{n^3\left(\ln n\cdots\ln_{q-1}n\right)^2}\right)\right) \\
&\quad + O\left(\frac{1}{\left(n\ln n\cdots\ln_q n\right)^3}\right) \\
&= \ln_{q+1}n - \frac{r}{n\ln n\cdots\ln_q n} - \frac{r^2}{2n^2\ln n\cdots\ln_q n} - \frac{r^2}{2(n\ln n)^2\ln_2 n\cdots\ln_q n} \\
&\quad - \cdots - \frac{r^2}{2\left(n\ln n\cdots\ln_q n\right)^2} - \frac{r^3(1+o(1))}{3n^3\ln n\cdots\ln_q n}.
\end{aligned}
\tag{3.48}
$$

Thus, formula (3.45) holds for $q+1$, too, which ends the proof. □

Lemma 3.6. *For fixed $r \in \mathbb{R}\setminus\{0\}$ and fixed $q\in\mathbb{N}$, the asymptotic representations*

$$
\begin{aligned}
\sqrt{\frac{\ln_q(n-r)}{\ln_q n}} &= 1 - \frac{r}{2n\ln n\cdots\ln_q n} - \frac{r^2}{4n^2\ln n\cdots\ln_q n} - \frac{r^2}{4(n\ln n)^2\ln_2 n\cdots\ln_q n} - \cdots \\
&\quad - \frac{r^2}{4\left(n\ln n\cdots\ln_{q-1}n\right)^2\ln_q n} - \frac{r^2}{8\left(n\ln n\cdots\ln_q n\right)^2} - \frac{r^3(1+o(1))}{6n^3\ln n\cdots\ln_q n},
\end{aligned}
\tag{3.49}
$$

$$
\sqrt{\frac{n-r}{n}} = 1 - \frac{r}{2n} - \frac{r^2}{8n^2} - \frac{r^3}{16n^3} + o\left(\frac{1}{n^3}\right)
\tag{3.50}
$$

hold for $n \to \infty$.

Proof. Both these relations are simple consequences of the asymptotic formula

$$
\sqrt{1-x} = 1 - \frac{1}{2}\, x - \frac{1}{8}\, x^2 - \frac{1}{16}\, x^3 + o\left(x^3\right) \quad \text{for } x \longrightarrow 0
\tag{3.51}
$$

and of Lemma 3.5 (for formula (3.49)). In the case of relation (3.49), we put

$$
x = \frac{r}{n\ln n\cdots\ln_q n} + \frac{r^2}{2n^2\ln n\cdots\ln_q n} + \cdots + \frac{r^2}{2\left(n\ln n\cdots\ln_{q-1}n\right)^2\ln_q n} + \frac{r^3(1+o(1))}{3n^3\ln n\cdots\ln_q n}
\tag{3.52}
$$

and in the case of relation (3.50), we put $x = r/n$. □

to Contents, page 1

Lemma 3.7. *For fixed $r \in \mathbb{R} \setminus \{0\}$ and fixed $q \in \mathbb{N}$, the asymptotic representation*

$$\sqrt{\frac{(n-r)}{n}\frac{\ln(n-r)}{\ln n}\cdots\frac{\ln_q(n-r)}{\ln_q n}} = 1 - r\left(\frac{1}{2n} + \frac{1}{2n\ln n} + \cdots + \frac{1}{2n\ln n\cdots\ln_q n}\right) - r^2\mu_q(n) - \frac{r^3}{16n^3} + o\left(\frac{1}{n^3}\right) \tag{3.53}$$

holds for $n \to \infty$.

Proof. We will prove relation (3.53) by induction with respect to q. For $q = 1$, (3.53) reduces to

$$\begin{aligned}\sqrt{\frac{(n-r)}{n}\frac{\ln(n-r)}{\ln n}} &= 1 - r\left(\frac{1}{2n} + \frac{1}{2n\ln n}\right) - r^2\mu_1(n) - \frac{r^3}{16n^3} + o\left(\frac{1}{n^3}\right)\\ &= 1 - r\left(\frac{1}{2n} + \frac{1}{2n\ln n}\right) - r^2\left(\frac{1}{8n^2} + \frac{1}{8(n\ln n)^2}\right) - \frac{r^3}{16n^3} + o\left(\frac{1}{n^3}\right).\end{aligned} \tag{3.54}$$

On the other hand, using Lemma 3.6, we get

$$\begin{aligned}&\sqrt{\frac{(n-r)}{n}\frac{\ln(n-r)}{\ln n}}\\ &= \left(1 - \frac{r}{2n} - \frac{r^2}{8n^2} - \frac{r^3}{16n^3} + o\left(\frac{1}{n^3}\right)\right)\\ &\quad\times\left(1 - \frac{r}{2n\ln n} - \frac{r^2}{4n^2\ln n} - \frac{r^2}{8(n\ln n)^2} - \frac{r^3(1+o(1))}{6n^3\ln n}\right)\\ &= 1 - \frac{r}{2n\ln n} - \frac{r^2}{4n^2\ln n} - \frac{r^2}{8(n\ln n)^2} - \frac{r}{2n} + \frac{r^2}{4n^2\ln n} - \frac{r^2}{8n^2} - \frac{r^3}{16n^3} + o\left(\frac{1}{n^3}\right)\\ &= 1 - r\left(\frac{1}{2n} + \frac{1}{2n\ln n}\right) - r^2\left(\frac{1}{8n^2} + \frac{1}{8(n\ln n)^2}\right) - \frac{r^3}{16n^3} + o\left(\frac{1}{n^3}\right).\end{aligned} \tag{3.55}$$

Thus, for $q = 1$, relation (3.53) holds. Now suppose that (3.53) holds for some q and prove that it holds for $q + 1$. In the following calculations, we use Lemma 3.6 and we skip some tedious expressions handling.

to Contents, page 1

$$\begin{aligned}
&\sqrt{\frac{(n-r)}{n}\frac{\ln(n-r)}{\ln n}\cdots\frac{\ln_{q+1}(n-r)}{\ln_{q+1}n}} \\
&\quad=\sqrt{\frac{(n-r)}{n}\frac{\ln(n-r)}{\ln n}\cdots\frac{\ln_{q}(n-r)}{\ln_{q}n}}\cdot\sqrt{\frac{\ln_{q+1}(n-r)}{\ln_{q+1}n}} \\
&\quad=\left(1-r\left(\frac{1}{2n}+\frac{1}{2n\ln n}+\cdots+\frac{1}{2n\ln n\cdots\ln_q n}\right)-r^2\mu_q(n)-\frac{r^3}{16n^3}+o\left(\frac{1}{n^3}\right)\right) \\
&\qquad\times\left(1-\frac{r}{2n\ln n\cdots\ln_{q+1}n}-\frac{r^2}{4n^2\ln n\cdots\ln_{q+1}n}-\cdots\right. \\
&\qquad\left.-\frac{r^2}{4(n\ln n\cdots\ln_q n)^2\ln_{q+1}n}-\frac{r^2}{8(n\ln n\cdots\ln_{q+1}n)^2}+o\left(\frac{1}{n^3}\right)\right) \\
&\quad=1-r\left(\frac{1}{2n}+\frac{1}{2n\ln n}+\cdots+\frac{1}{2n\ln n\cdots\ln_{q+1}n}\right)-r^2\mu_{q+1}(n)-\frac{r^3}{16n^3}+o\left(\frac{1}{n^3}\right).
\end{aligned}\tag{3.56}$$

We can see that formula (3.53) holds for $q+1$, too, which ends the proof. □

Now we are ready to prove that there exists a bounded positive solution of (3.38). Remind that functions p_ℓ and v_ℓ were defined by (3.40) and (3.41), respectively.

Theorem 3.8. *Let $\omega : \mathbb{Z}_a^\infty \to \mathbb{R}$ satisfy the inequality*

$$|\omega(n)| \le \varepsilon\left(\frac{k}{k+1}\right)^k\cdot\frac{k(2k^2+k-1)}{16n^3(k+1)}, \quad n\in\mathbb{Z}_a^\infty, \tag{3.57}$$

for a fixed $\varepsilon\in(0,1)$. Suppose that there exists an integer $\ell\ge 0$ such that the function p satisfies the inequalities

$$0 < p(n) \le p_\ell(n) + \omega(n) \tag{3.58}$$

for every $n\in\mathbb{Z}_a^\infty$. Then there exists a solution $v=v^*(n)$, $n\in\mathbb{Z}_{a-k}^\infty$ of (3.38) such that for n sufficiently large the inequalities

$$0 < v^*(n) < v_\ell(n) \tag{3.59}$$

hold.

Proof. Show that all the assumptions of Theorem 3.1 are fulfilled. For (3.38), $f(n, v_1, \ldots, v_{k+1}) = -p(n)v_{k+1}$. This is a continuous function. Put

$$b(n) := 0, \qquad c(n) := v_\ell(n). \tag{3.60}$$

to Contents, page 1

We have to prove that for every $v_2, \ldots, v_{k+1}$ such that $b(n-i+1) < v_i < c(n-i+1)$, $i = 2, \ldots, k+1$, the inequalities (3.4) and (3.5) hold for n sufficiently large. Start with (3.4). That gives that for $b(n-k) < v_{k+1} < c(n-k)$, it has to be

$$0 - p(n) \cdot v_{k+1} < 0. \tag{3.61}$$

This certainly holds, because the function p is positive and so is v_{k+1}.

Next, according to (3.5), we have to prove that

$$\nu_\ell(n) - p(n) v_{k+1} > \nu_\ell(n+1) \tag{3.62}$$

which is equivalent to the inequality

$$-p(n) v_{k+1} > \nu_\ell(n+1) - \nu_\ell(n). \tag{3.63}$$

Denote the left-hand side of (3.63) as $L_{(3.63)}$. As $v_{k+1} < c(n-k) = \nu_\ell(n-k)$ and as by (3.40), (3.58), and (3.57)

$$p(n) \leq \left(\frac{k}{k+1}\right)^k \cdot \left(\frac{1}{k+1} + k\mu_\ell(n)\right) + \varepsilon \left(\frac{k}{k+1}\right)^k \cdot \frac{k(2k^2+k-1)}{16n^3(k+1)}, \tag{3.64}$$

we have

$$\begin{aligned} L_{(3.63)} &> -\left(\frac{k}{k+1}\right)^k \left(\frac{1}{k+1} + k\mu_\ell(n) + \varepsilon \cdot \frac{k(2k^2+k-1)}{16n^3(k+1)}\right) \\ &\quad \times \left(\frac{k}{k+1}\right)^{n-k} \sqrt{(n-k)\ln(n-k)\cdots\ln_\ell(n-k)} \\ &= -\left(\frac{k}{k+1}\right)^n \left(\frac{1}{k+1} + k\mu_\ell(n) + \varepsilon \cdot \frac{k(2k^2+k-1)}{16n^3(k+1)}\right) \cdot \sqrt{(n-k)\ln(n-k)\cdots\ln_\ell(n-k)}. \end{aligned} \tag{3.65}$$

Further, we can easily see that

$$\nu_\ell(n+1) - \nu_\ell(n) = \left(\frac{k}{k+1}\right)^n \sqrt{n \ln n \cdots \ln_\ell n} \left(\frac{k}{k+1} \sqrt{\frac{(n+1)}{n} \frac{\ln(n+1)}{\ln n} \cdots \frac{\ln_\ell(n+1)}{\ln_\ell n}} - 1\right). \tag{3.66}$$

to Contents, page 1

Thus, to prove (3.63), it suffices to show that for n sufficiently large,

$$\begin{aligned}
&-\left(\frac{1}{k+1}+k\mu_\ell(n)+\varepsilon\cdot\frac{k(2k^2+k-1)}{16n^3(k+1)}\right)\sqrt{\frac{(n-k)}{n}\frac{\ln(n-k)}{\ln n}\cdots\frac{\ln_\ell(n-k)}{\ln_\ell n}}\\
&\quad>\frac{k}{k+1}\sqrt{\frac{(n+1)}{n}\frac{\ln(n+1)}{\ln n}\cdots\frac{\ln_\ell(n+1)}{\ln_\ell n}}-1.
\end{aligned}\tag{3.67}$$

Denote the left-hand side of inequality (3.67) as $L_{(3.67)}$ and the right-hand side as $R_{(3.67)}$. Using Lemma 3.7 with $r=k$ and $q=\ell$, we can write

$$\begin{aligned}
L_{(3.67)}&=-\left(\frac{1}{k+1}+k\mu_\ell(n)+\varepsilon\cdot\frac{k(2k^2+k-1)}{16n^3(k+1)}\right)\\
&\quad\times\left(1-k\left(\frac{1}{2n}+\frac{1}{2n\ln n}+\cdots+\frac{1}{2n\ln n\cdots\ln_\ell n}\right)-k^2\mu_\ell(n)-\frac{k^3}{16n^3}+o\left(\frac{1}{n^3}\right)\right)\\
&=-\frac{1}{k+1}+\frac{k}{k+1}\left(\frac{1}{2n}+\frac{1}{2n\ln n}+\cdots+\frac{1}{2n\ln n\cdots\ln_\ell n}\right)\\
&\quad+\frac{k^2}{k+1}\mu_\ell(n)+\frac{k^3}{16n^3(k+1)}-k\mu_\ell(n)+\frac{k^2}{16n^3}-\varepsilon\cdot\frac{k(2k^2+k-1)}{16n^3(k+1)}+o\left(\frac{1}{n^3}\right)\\
&=-\frac{1}{k+1}+\frac{k}{k+1}\left(\frac{1}{2n}+\frac{1}{2n\ln n}+\cdots+\frac{1}{2n\ln n\cdots\ln_\ell n}\right)\\
&\quad-\frac{k}{k+1}\mu_\ell(n)+\frac{2k^3(1-\varepsilon)+k^2(1-\varepsilon)+k\varepsilon}{16n^3(k+1)}+o\left(\frac{1}{n^3}\right).
\end{aligned}\tag{3.68}$$

Using Lemma 3.7 with $r=-1$ and $q=\ell$, we get for $R_{(3.67)}$

$$\begin{aligned}
R_{(3.67)}&=\frac{k}{k+1}\left(1+\frac{1}{2n}+\frac{1}{2n\ln n}+\cdots+\frac{1}{2n\ln n\cdots\ln_\ell n}-\mu_\ell(n)+\frac{1}{16n^3}+o\left(\frac{1}{n^3}\right)\right)-1\\
&=\frac{-1}{k+1}+\frac{k}{k+1}\left(\frac{1}{2n}+\frac{1}{2n\ln n}+\cdots+\frac{1}{2n\ln n\cdots\ln_\ell n}\right)\\
&\quad-\frac{k}{k+1}\cdot\mu_\ell(n)+\frac{k}{16n^3(k+1)}+o\left(\frac{1}{n^3}\right).
\end{aligned}\tag{3.69}$$

It is easy to see that the inequality (3.67) reduces to

$$\frac{2k^3(1-\varepsilon)+k^2(1-\varepsilon)+k\varepsilon}{16n^3(k+1)}+o\left(\frac{1}{n^3}\right)>\frac{k}{16n^3(k+1)}+o\left(\frac{1}{n^3}\right).\tag{3.70}$$

to Contents, page 1

This inequality is equivalent to

$$\frac{k\left(2k^2(1-\varepsilon)+k(1-\varepsilon)-(1-\varepsilon)\right)}{16n^3(k+1)}+o\left(\frac{1}{n^3}\right)>0. \tag{3.71}$$

The last inequality holds for n sufficiently large because $k \geq 1$ and $1-\varepsilon \in (0,1)$. We have proved that all the assumptions of Theorem 3.1 are fulfilled and hence there exists a solution of (3.38) satisfying conditions (3.59). □

Acknowledgment

This paper was supported by Grant P201/10/1032 of the Czech Grant Agency (Prague), by the Council of Czech Government MSM 00216 30529, and by the Grant FEKT-S-11-2-921 of Faculty of Electrical Engineering and Communication, Brno University of Technology.

References

[1] J. Diblík, "Asymptotic behaviour of solutions of systems of discrete equations via Liapunov type technique," *Computers and Mathematics with Applications*, vol. 45, no. 6–9, pp. 1041–1057, 2003, Advances in difference equations, I.

[2] J. Diblík, "Discrete retract principle for systems of discrete equations," *Computers and Mathematics with Applications*, vol. 42, no. 3–5, pp. 515–528, 2001, Advances in difference equations, II.

[3] J. Diblík, "Anti-Lyapunov method for systems of discrete equations," *Nonlinear Analysis*, vol. 57, no. 7–8, pp. 1043–1057, 2004.

[4] J. Diblík, M. Migda, and E. Schmeidel, "Bounded solutions of nonlinear discrete equations," *Nonlinear Analysis*, vol. 65, no. 4, pp. 845–853, 2006.

[5] J. Diblík, I. Růžičková, and M. Růžičková, "A general version of the retract method for discrete equations," *Acta Mathematica Sinica*, vol. 23, no. 2, pp. 341–348, 2007.

[6] J. Diblík and I. Růžičková, "Compulsory asymptotic behavior of solutions of two-dimensional systems of difference equations," in *Proceedings of the 9th International Conference on Difference Equations and Discrete Dynamical Systems*, pp. 35–49, World Scientific Publishing, University of Southern California, Los Angeles, Calif, USA, 2005.

[7] J. Diblík and I. Hlavičková, "Asymptotic properties of solutions of the discrete analogue of the Emden-Fowler equation," in *Advances in Discrete Dynamical Systems*, vol. 53 of *Advanced Studies in Pure Mathematics*, pp. 23–32, Mathematical Society of Japan, Tokyo, Japan, 2009.

[8] G. Ladas, C. G. Philos, and Y. G. Sficas, "Sharp conditions for the oscillation of delay difference equations," *Journal of Applied Mathematics and Simulation*, vol. 2, no. 2, pp. 101–111, 1989.

[9] I. Györi and G. Ladas, *Oscillation Theory of Delay Differential Equations*, Clarendon Press, Alderley, UK, 1991.

[10] J. Baštinec and J. Diblík, "Subdominant positive solutions of the discrete equation $\Delta u(k+n) = -p(k)u(k)$," *Abstract and Applied Analysis*, no. 6, pp. 461–470, 2004.

[11] J. Baštinec and J. Diblík, "Remark on positive solutions of discrete equation $\Delta u(k+n) = -p(k)u(k)$," *Nonlinear Analysis*, vol. 63, no. 5-7, pp. e2145–e2151, 2004.

[12] J. Baštinec, J. Diblík, and B. Zhang, "Existence of bounded solutions of discrete delayed equations," in *Proceedings of the Sixth International Conference on Difference Equations and Applications*, pp. 359–366, CRC, Boca Raton, Fla, USA, 2004.

[13] R. P. Agarwal, M. Bohner, and W.-T. Li, *Nonoscillation and Oscillation: Theory for Functional Differential Equations*, Marcel Dekker, New York, NY, USA, 2004.

[14] L. Berezansky and E. Braverman, "On existence of positive solutions for linear difference equations with several delays," *Advances in Dynamical Systems and Applications*, vol. 1, no. 1, pp. 29–47, 2006.

[15] G. E. Chatzarakis and I. P. Stavroulakis, "Oscillations of first order linear delay difference equations," *The Australian Journal of Mathematical Analysis and Applications*, vol. 3, no. 1, article 14, 2006.

to Contents, page 1

[16] D. Chengjun and S. Qiankun, "Boundedness and stability for discrete-time delayed neural network with complex-valued linear threshold neurons," *Discrete Dynamics in Nature and Society*, vol. 2010, article 368379, 2010.

[17] J. Čermák, "Asymptotic bounds for linear difference systems," *Advances in Difference Equations*, vol. 2010, article 182696, 2010.

[18] I. Györi and M. Pituk, "Asymptotic formulae for the solutions of a linear delay difference equation," *Journal of Mathematical Analysis and Applications*, vol. 195, no. 2, pp. 376–392, 1995.

[19] P. Karajani and I. P. Stavroulakis, "Oscillation criteria for second-order delay and functional equations," *Studies of the University of Žilina. Mathematical Series*, vol. 18, no. 1, pp. 17–26, 2004.

[20] L. K. Kikina and I. P. Stavroulakis, "A survey on the oscillation of solutions of first order delay difference equations," *CUBO, A Mathematical Journal*, vol. 7, no. 2, pp. 223–236, 2005.

[21] L. K. Kikina and I. P. Stavroulakis, "Oscillation criteria for second-order delay, difference, and functional equations," *International Journal of Differential Equations*, vol. 2010, article 598068, 2010.

[22] M. Kipnis and D. Komissarova, "Stability of a delay difference system," *Advances in Difference Equations*, vol. 2006, article 31409, 2006.

[23] E. Liz, "Local stability implies global stability in some one-dimensional discrete single-species models," *Discrete and Continuous Dynamical Systems. Series B*, vol. 7, no. 1, pp. 191–199, 2007.

[24] R. Medina and M. Pituk, "Nonoscillatory solutions of a second-order difference equation of Poincaré type," *Applied Mathematics Letters*, vol. 22, no. 5, pp. 679–683, 2009.

[25] I. P. Stavroulakis, "Oscillation criteria for first order delay difference equations," *Mediterranean Journal of Mathematics*, vol. 1, no. 2, pp. 231–240, 2004.

[26] J. Baštinec, J. Diblík, and Z. Šmarda, "Existence of positive solutions of discrete linear equations with a single delay," *Journal of Difference Equations and Applications*, vol. 16, no. 9, pp. 1047–1056, 2010.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 986343, 12 pages
doi:10.1155/2011/986343

Research Article

Asymptotic Behavior of Solutions to Half-Linear q-Difference Equations

Pavel Řehák

Institute of Mathematics, Academy of Sciences of the Czech Republic, Žižkova 22, 61662 Brno, Czech Republic

Correspondence should be addressed to Pavel Řehák, rehak@math.cas.cz

Received 11 October 2010; Accepted 18 November 2010

Academic Editor: Elena Braverman

We derive necessary and sufficient conditions for (some or all) positive solutions of the half-linear q-difference equation $D_q(\Phi(D_q y(t))) + p(t)\Phi(y(qt)) = 0$, $t \in \{q^k : k \in \mathbb{N}_0\}$ with $q > 1$, $\Phi(u) = |u|^{\alpha-1} \operatorname{sgn} u$ with $\alpha > 1$, to behave like q-regularly varying or q-rapidly varying or q-regularly bounded functions (that is, the functions y, for which a special limit behavior of $y(qt)/y(t)$ as $t \to \infty$ is prescribed). A thorough discussion on such an asymptotic behavior of solutions is provided. Related Kneser type criteria are presented.

1. Introduction

In this paper we recall and survey the theory of q-Karamata functions, that is, of the functions $y : q^{\mathbb{N}_0} \to (0,\infty)$, where $q^{\mathbb{N}_0} := \{q^k : k \in \mathbb{N}_0\}$ with $q > 1$, and for which some special limit behavior of $y(qt)/y(t)$ as $t \to \infty$ is prescribed, see [1–3]. This theory corresponds with the classical "continuous" theory of regular variation, see, for example, [4], but shows some special features (see Section 2), not known in the continuous case, which are due to the special structure of $q^{\mathbb{N}_0}$. The theory of q-Karamata functions provides a powerful tool, which we use in this paper to establish sufficient and necessary conditions for some or all positive solutions of the half-linear q-difference equation

$$D_q\big(\Phi(D_q y(t))\big) + p(t)\Phi\big(y(qt)\big) = 0, \tag{1.1}$$

where $\Phi(u) = |u|^{\alpha-1} \operatorname{sgn} u$ with $\alpha > 1$, to behave like q-regularly varying or q-rapidly varying or q-regularly bounded functions. We stress that there is no sign condition on p. We also

to Contents, page 1

present Kneser type (non)oscillation criteria for (1.1), existing as well as new ones, which are somehow related to our asymptotic results.

The main results of this paper can be understood as a q-version of the continuous results for

$$\left(\Phi\left(y'(t)\right)\right)' + p(t)\Phi\left(y(t)\right) = 0 \tag{1.2}$$

from [5] (with noting that some substantial differences between the parallel results are revealed), or as a half-linear extension of the results for $D_q^2 y(t) + p(t)y(qt) = 0$ from [1]. In addition, we provide a thorough description of asymptotic behavior of solutions to (1.1) with respect to the limit behavior of $t^\alpha p(t)$ in the framework of q-Karamata theory. For an explanation why the q-Karamata theory and its applications are not included in a general theory of regular variation on measure chains see [6]. For more information on (1.2) see, for example, [7]. Many applications of the theory of regular variation in differential equations can be found, for example, in [8]. Linear q-difference equations were studied, for example, in [1, 9–11]; for related topics see, for example, [12, 13]. Finally note that the theory of q-calculus is very extensive with many aspects; some people speak bout different tongues of q-calculus. In our paper we follow essentially its "time-scale dialect".

2. Preliminaries

We start with recalling some basic facts about q-calculus. For material on this topic see [9, 12, 13]. See also [14] for the calculus on time-scales which somehow contains q-calculus. First note that some of the below concepts may appear to be described in a "nonclassical q-way", see, for example, our definition of q-integral versus original Jackson's definition [9, 12, 13], or the q-exponential function. But, working on the lattice $q^{\mathbb{N}_0}$ (which is a time-scale), we can introduce these concepts in an alternative and "easier" way (and, basically, we avoid some classical q-symbols). Our definitions, of course, naturally correspond with the original definitions. The q-derivative of a function $f : q^{\mathbb{N}_0} \to \mathbb{R}$ is defined by $D_q f(t) = [f(qt) - f(t)]/[(q-1)t]$. The q-integral $\int_a^b f(t)\mathrm{d}_q t$, $a, b \in q^{\mathbb{N}_0}$, of a function $f : q^{\mathbb{N}_0} \to \mathbb{R}$ is defined by $\int_a^b f(t)\mathrm{d}_q t = (q-1)\sum_{t\in[a,b)\cap q^{\mathbb{N}_0}} t f(t)$ if $a < b$; $\int_a^b f(t)\mathrm{d}_q t = 0$ if $a = b$; $\int_a^b f(t)\mathrm{d}_q t = (1-q)\sum_{t\in[b,a)\cap q^{\mathbb{N}_0}} t f(t)$ if $a > b$. The improper q-integral is defined by $\int_a^\infty f(t)\mathrm{d}_q t = \lim_{b\to\infty}\int_a^b f(t)\mathrm{d}_q t$. We use the notation $[a]_q = (q^a - 1)/(q-1)$ for $a \in \mathbb{R}$. Note that $\lim_{q\to 1^+}[a]_q = a$. It holds that $D_q t^\vartheta = [\vartheta]_q t^{\vartheta-1}$. In view of the definition of $[a]_q$, it is natural to introduce the notation $[\infty]_q = \infty$, $[-\infty]_q = 1/(1-q)$. For $p \in \mathcal{R}$ (i.e., for $p : q^{\mathbb{N}_0} \to \mathbb{R}$ satisfying $1 + (q-1)tp(t) \neq 0$ for all $t \in q^{\mathbb{N}_0}$) we denote $e_p(t,s) = \prod_{u\in[s,t)\cap q^{\mathbb{N}_0}}[(q-1)up(u) + 1]$ for $s < t$, $e_p(t,s) = 1/e_p(s,t)$ for $s > t$, and $e_p(t,t) = 1$, where $s, t \in q^{\mathbb{N}_0}$. For $p \in \mathcal{R}$, $e(\cdot, a)$ is a solution of the IVP $D_q y = p(t)y$, $y(a) = 1$, $t \in q^{\mathbb{N}_0}$. If $s \in q^{\mathbb{N}_0}$ and $p \in \mathcal{R}^+$, where $\mathcal{R}^+ = \{p \in \mathcal{R} : 1 + (q-1)tp(t) > 0 \text{ for all } t \in q^{\mathbb{N}_0}\}$. then $e_p(t,s) > 0$ for all $t \in q^{\mathbb{N}_0}$. If $p, r \in \mathcal{R}$, then $e_p(t,s)e_p(s,u) = e_p(t,u)$ and $e_p(t,s)e_r(t,s) = e_{p+r+t(q-1)pr}(t,s)$. Intervals having the subscript q denote the intervals in $q^{\mathbb{N}_0}$, for example, $[a,\infty)_q = \{a, aq, aq^2, \ldots\}$ with $a \in q^{\mathbb{N}_0}$.

to Contents, page 1

Next we present auxiliary statements which play important roles in proving the main results. Define $F : (0,\infty) \to \mathbb{R}$ by $F(x) = \Phi(x/q - 1/q) - \Phi(1 - 1/x)$ and $h : (\Phi([-\infty]_q), \infty) \to \mathbb{R}$ by

$$h(x) = \frac{x}{1-q^{1-\alpha}}\left[1 - \left(1 + (q-1)\Phi^{-1}(x)\right)^{1-\alpha}\right]. \tag{2.1}$$

For $y : q^{\mathbb{N}_0} \to \mathbb{R} \setminus \{0\}$ define the operator $\mathcal{L}$ by

$$\mathcal{L}[y](t) = \Phi\left(\frac{y(q^2t)}{qy(qt)} - \frac{1}{q}\right) - \Phi\left(1 - \frac{y(t)}{y(qt)}\right). \tag{2.2}$$

We denote $\omega_q = ([(\alpha-1)/\alpha]_q)^{\alpha}$. Let β mean the conjugate number of α, that is, $1/\alpha + 1/\beta = 1$.

The following lemma lists some important properties of F, h, $\mathcal{L}$ and relations among them.

Lemma 2.1. (i) *The function F has the global minimum on $(0,\infty)$ at $q^{(\alpha-1)/\alpha}$ with*

$$F\left(q^{(\alpha-1)/\alpha}\right) = -\frac{\omega_q(q-1)^{\alpha}}{q^{\alpha-1}} \tag{2.3}$$

and $F(1) = 0 = F(q)$. Further, F is strictly decreasing on $(0, q^{(\alpha-1)/\alpha})$ and strictly increasing on $(q^{(\alpha-1)/\alpha}, \infty)$ with $\lim_{x\to 0+} F(x) = \infty$, $\lim_{t\to\infty} F(x) = \infty$.

(ii) *The graph of $x \mapsto h(x)$ is a parabola-like curve with the minimum at the origin. The graph of $x \mapsto h(x) + \gamma_\alpha$ touches the line $x \mapsto x$ at $x = \lambda_0 := ([(\alpha-1)/\alpha]_q)^{\alpha-1}$. The equation $h(\lambda) + \gamma = \lambda$ has*

(a) *no real roots if $\gamma > \omega_q/[\alpha-1]_q$,*

(b) *the only root λ_0 if $\gamma = \omega_q/[\alpha-1]_q$,*

(c) *two real roots λ_1, λ_2 with $0 < \lambda_1 < \lambda_0 < \lambda_2 < 1$ if $\gamma \in (0, \omega_q/[\alpha-1]_q)$,*

(d) *two real roots 0 and 1 if $\gamma = 0$,*

(e) *two real roots λ_1, λ_2 with $\lambda_1 < 0 < 1 < \lambda_2$ if $\gamma < 0$.*

(iii) *It holds that $F(q^{\vartheta_1}) = F(q^{\vartheta_2})$, where $\vartheta_i = \log_q[(q-1)\Phi^{-1}(\lambda_i) + 1]$, $i = 1, 2$, with $\lambda_1 < \lambda_2$ being the real roots of the equation $\lambda = h(\lambda) + A$ with $A \in (-\infty, \omega_q/[\alpha-1]_q)$.*

(iv) *If $q \to 1+$, then $h(x) \to |x|^{\beta}$.*

(v) *For $\vartheta \in \mathbb{R}$ it hold that $\Phi([\vartheta]_q)[1-\vartheta]_{q^{\alpha-1}} = \Phi([\vartheta]_q) - h(\Phi([\vartheta]_q))$.*

(vi) *For $\vartheta \in \mathbb{R}$ it hold that $F(q^{\vartheta}) = (q-1)^{\alpha}[1-\alpha]_q\Phi([\vartheta]_q)[1-\vartheta]_{q^{\alpha-1}}$.*

(vii) *For $y \neq 0$, (1.1) can be written as $\mathcal{L}[y](t) = -(q-1)^{\alpha}t^{\alpha}p(t)$.*

(viii) *If the $\lim_{t\to\infty} y(qt)/y(t)$ exists as a positive real number, then $\lim_{t\to\infty}\mathcal{L}[y](t) = \lim_{t\to\infty} F(y(qt)/y(t))$.*

Proof. We prove only (iii). The proofs of other statements are either easy or can be found in [3].

to Contents, page 1

(iii) Let λ_1, λ_2 be the real roots of $\lambda = h(\lambda) + A$. We have $\lambda_i = \Phi([\vartheta_i]_q)$, $i = 1, 2$, and so, by virtue of identities (v) and (vi), we get $F(q^{\vartheta_1}) = (q-1)^\alpha [1-\alpha]_q (\lambda_1 - h(\lambda_1)) = (q-1)^\alpha [1-\alpha]_q A = (q-1)^\alpha [1-\alpha]_q (\lambda_2 - h(\lambda_2)) = F(q^{\vartheta_2})$. □

Next we define the basic concepts of q-Karamata theory. Note that the original definitions (see [1–3]) was more complicated; they were motivated by the classical continuous and the discrete (on the uniform lattices) theories. But soon it has turned out that simpler (and equivalent)definitions can be established. Also, there is no need to introduce the concept of normality, since every q-regularly varying or q-rapidly varying or q-regularly bounded function is automatically normalized. Such (and some other) simplifications are not possible in the original continuous theory or in the classical discrete theory; in q-calculus, they are practicable thanks to the special structure of $q^{\mathbb{N}_0}$, which is somehow natural for examining regularly varying behavior.

For $f : q^{\mathbb{N}_0} \to (0, \infty)$ denote

$$K_* = \liminf_{t\to\infty} \frac{f(qt)}{f(t)}, \qquad K^* = \limsup_{t\to\infty} \frac{f(qt)}{f(t)}, \qquad K = \lim_{t\to\infty} \frac{f(qt)}{f(t)}. \tag{2.4}$$

Definition 2.2. A function $f : q^{\mathbb{N}_0} \to (0, \infty)$ is said to be

(i) *q-regularly varying of index ϑ*, $\vartheta \in \mathbb{R}$, if $K = q^\vartheta$; we write $f \in \mathcal{RV}_q(\vartheta)$,

(ii) *q-slowly varying* if $K = 1$; we write $f \in \mathcal{SV}_q$,

(iii) *q-rapidly varying of index ∞* if $K = \infty$; we write $f \in \mathcal{RPV}_q(\infty)$,

(iv) *q-rapidly varying of index $-\infty$* if $K = 0$; we write $f \in \mathcal{RPV}_q(-\infty)$,

(v) *q-regularly bounded* if $0 < K_* \le K^* < \infty$; we write $f \in \mathcal{RB}_q$.

Clearly, $\mathcal{SV}_q = \mathcal{RV}_q(0)$. We have defined q-regular variation, q-rapid variation, and q-regular boundedness at infinity. If we consider a function $f : q^{\mathbb{Z}} \to (0, \infty)$, $q^{\mathbb{Z}} := \{q^k : k \in \mathbb{Z}\}$, then $f(t)$ is said to be q-regularly varying/q-rapidly varying/q-regularly bounded at zero if $f(1/t)$ is q-regularly varying/q-rapidly varying/q-regularly bounded at infinity. But it is apparent that it is sufficient to examine just the behavior at ∞.

Next we list some selected important properties of the above-defined functions. We define $\tau : [1, \infty) \to q^{\mathbb{N}_0}$ as $\tau(x) = \max\{s \in q^{\mathbb{N}_0} : s \le x\}$.

Proposition 2.3. (i) $f \in \mathcal{RV}_q(\vartheta) \Leftrightarrow \lim_{t\to\infty} tD_q f(t)/f(t) = [\vartheta]_q$.

(ii) $f \in \mathcal{RV}_q(\vartheta) \Leftrightarrow f(t) = \varphi(t) e_\psi(t, 1)$, *where a positive* φ *satisfies* $\lim_{t\to\infty} \varphi(t) = C \in (0, \infty)$, $\lim_{t\to\infty} t\psi(t) = [\vartheta]_q$, $\psi \in \mathcal{R}^+$ *(w.l.o.g., φ can be replaced by C).*

(iii) $f \in \mathcal{RV}_q(\vartheta) \Leftrightarrow f(t) = t^\vartheta L(t)$, *where* $L \in \mathcal{SV}_q$.

(iv) $f \in \mathcal{RV}_q(\vartheta) \Leftrightarrow f(t)/t^\gamma$ *is eventually increasing for each* $\gamma < \vartheta$ *and* $f(t)/t^\eta$ *is eventually decreasing for each* $\eta > \vartheta$.

(v) $f \in \mathcal{RV}_q(\vartheta) \Leftrightarrow \lim_{t\to\infty} f(\tau(\lambda t))/f(t) = (\tau(\lambda))^\vartheta$ *for every* $\lambda \ge 1$.

(vi) $f \in \mathcal{RV}_q(\vartheta) \Leftrightarrow R : [1, \infty) \to (0, \infty)$ *defined by* $R(x) = f(\tau(x))(x/\tau(x))^\vartheta$ *for* $x \in [1, \infty)$ *is regularly varying of index* ϑ.

(vii) $f \in \mathcal{RV}_q(\vartheta) \Rightarrow \lim_{t\to\infty} \log f(t)/\log t = \vartheta$.

Proof. See [2]. □

to Contents, page 1

Proposition 2.4. (i) $f \in \mathcal{RPV}_q(\pm\infty) \Leftrightarrow \lim_{t\to\infty} tD_q f(t)/f(t) = [\pm\infty]_q$.

(ii) *$f \in \mathcal{RPV}_q(\pm\infty) \Leftrightarrow f(t) = \varphi(t)e_\psi(t,1)$, where a positive φ satisfies $\liminf_{t\to\infty}\varphi(qt)/\varphi(t) > 0$ for index ∞, $\limsup_{t\to\infty}\varphi(qt)/\varphi(t) < \infty$ for index $-\infty$, and $\lim_{t\to\infty} t\psi(t) = [\pm\infty]_q$, $\psi \in \mathcal{R}^+$ (w.l.o.g., φ can be replaced by $C \in (0,\infty)$).*

(iii) *$f \in \mathcal{RPV}_q(\pm\infty) \Leftrightarrow$ for each $\vartheta \in [0,\infty)$, $f(t)/t^\vartheta$ is eventually increasing (towards ∞) for index ∞ and $f(t)t^\vartheta$ is eventually decreasing (towards 0) for index $-\infty$.*

(iv) *$f \in \mathcal{RPV}_q(\pm\infty) \Leftrightarrow$ for every $\lambda \in [q,\infty)$ it holds, $\lim_{t\to\infty} f(\tau(\lambda t))/f(t) = \infty$ for index ∞ and $\lim_{t\to\infty} f(\tau(\lambda t))/f(t) = 0$ for index $-\infty$.*

(v) *Let $R : [1,\infty) \to (0,\infty)$ be defined by $R(x) = f(\tau(x))$ for $x \in [1,\infty)$. If R is rapidly varying of index $\pm\infty$, then $f \in \mathcal{RPV}_q(\pm\infty)$. Conversely, if $f \in \mathcal{RPV}_q(\pm\infty)$, then $\lim_{x\to\infty} R(\lambda x)/R(x) = \infty$, resp., $\lim_{x\to\infty} R(\lambda x)/R(x) = 0$ for $\lambda \in [q,\infty)$.*

(vi) *$f \in \mathcal{RPV}_q(\pm\infty) \Rightarrow \lim_{t\to\infty} \log f(t)/\log t = \pm\infty$.*

Proof. We prove only the "if" part of (iii). The proofs of (iv), (v), and (vi) can be found in [1]. The proofs of other statements can be found in [3].

Assume that $f(t)/t^\vartheta$ is eventually increasing (towards ∞) for each $\vartheta \in [0,\infty)$. Because of monotonicity, we have $f(t)/t^\vartheta \le f(qt)/(q^\vartheta t^\vartheta)$, and so $f(qt)/f(t) \ge q^\vartheta$ for large t. Since ϑ is arbitrary, we have $f(qt)/f(t) \to \infty$ as $t \to \infty$, thus $f \in \mathcal{RPV}_q(\infty)$. The case of the index $-\infty$ can be treated in a similar way. □

Proposition 2.5. (i) $f \in \mathcal{RB}_q \Leftrightarrow [-\infty]_q < \liminf_{t\to\infty} tD_q f(t)/f(t) \le \limsup_{t\to\infty} tD_q f(t)/f(t) < [\infty]_q$.

(ii) *$f \in \mathcal{RB}_q \Leftrightarrow f(t) = t^\vartheta\varphi(t)e_\psi(t,1)$, where $0 < C_1 \le \varphi(t) \le C_2 < \infty$, $[-\infty]_q < D_1 \le t\psi(t) \le D_2 < [\infty]_q$ (w.l.o.g., φ can be replaced by $C \in (0,\infty)$).*

(iii) *$f \in \mathcal{RB}_q \Leftrightarrow f(t)/t^{\gamma_1}$ is eventually increasing and $f(t)/t^{\gamma_2}$ is eventually decreasing for some $\gamma_1 < \gamma_2$ (w.l.o.g., monotonicity can be replaced by almost monotonicity; a function $f : q^{\mathbb{N}_0} \to (0,\infty)$ is said to be almost increasing (almost decreasing) if there exists an increasing (decreasing) function $g : q^{\mathbb{N}_0} \to (0,\infty)$ and $C, D \in (0,\infty)$ such that $Cg(t) \le f(t) \le Dg(t)$).*

(iv) *$f \in \mathcal{RB}_q \Leftrightarrow 0 < \liminf_{t\to\infty} f(\tau(\lambda t))/f(t) \le \limsup_{t\to\infty} f(\tau(\lambda t))/f(t) < \infty$ for every $\lambda \in [q,\infty)$ or for every $\lambda \in (0,1)$.*

(v) *$f \in \mathcal{RB}_q \Leftrightarrow R : [1,\infty) \to (0,\infty)$ defined by $R(x) = f(\tau(x))$ for $x \in [1,\infty)$ is regularly bounded.*

(vi) *$f \in \mathcal{RB}_q \Rightarrow -\infty < \liminf_{t\to\infty} \log f(t)/\log t \le \limsup_{t\to\infty} \log f(t)/\log t < \infty$.*

Proof. See [1]. □

For more information on q-Karamata theory see [1–3].

3. Asymptotic Behavior of Solutions to (1.1) in the Framework of q-Karamata Theory

First we establish necessary and sufficient conditions for positive solutions of (1.1) to be q-regularly varying or q-rapidly varying or q-regularly bounded. Then we use this result to provide a thorough discussion on Karamata-like behavior of solutions to (1.1).

Theorem 3.1. (i) *Equation* (1.1) *has eventually positive solutions u, v such that $u \in \mathcal{RV}_q(\vartheta_1)$ and $v \in \mathcal{RV}_q(\vartheta_2)$ if and only if*

$$\lim_{t\to\infty} t^\alpha p(t) = P \in \left(-\infty, \frac{\omega_q}{q^{\alpha-1}}\right), \tag{3.1}$$

to Contents, page 1

where $\vartheta_i = \log_q[(q-1)\Phi^{-1}(\lambda_i)+1]$, $i = 1,2$, *with* $\lambda_1 < \lambda_2$ *being the real roots of the equation* $\lambda = h(\lambda) - P/[1-\alpha]_q$. *For the indices* ϑ_i, $i = 1,2$, *it holds that* $\vartheta_1 < 0 < 1 < \vartheta_2$ *provided* $P < 0$; $\vartheta_1 = 0$, $\vartheta_2 = 1$ *provided* $P = 0$; $0 < \vartheta_1 < (\alpha-1)/\alpha < \vartheta_2 < 1$ *provided* $P > 0$. *Any of two conditions* $u \in \mathcal{RV}_q(\vartheta_1)$ *and* $v \in \mathcal{RV}_q(\vartheta_2)$ *implies* (3.1).

(ii) *Let* (1.1) *be nonoscillatory (which can be guaranteed, for example, by* $t^\alpha p(t) \le \omega_q/q^{\alpha-1}$ *for large* t*; with the note that it allows* (3.2)*). Equation* (1.1) *has an eventually positive solution* u *such that* $u \in \mathcal{RV}_q((\alpha-1)/\alpha)$ *if and only if*

$$\lim_{t\to\infty} t^\alpha p(t) = \frac{\omega_q}{q^{\alpha-1}}. \tag{3.2}$$

All eventually positive solutions of (1.1) *are* q*-regularly varying of index* $(\alpha-1)/\alpha$ *provided* (3.2) *holds.*

(iii) *Equation* (1.1) *has eventually positive solutions* u, v *such that* $u \in \mathcal{RPV}_q(-\infty)$ *and* $u \in \mathcal{RPV}_q(\infty)$ *if and only if*

$$\lim_{t\to\infty} t^\alpha p(t) = -\infty. \tag{3.3}$$

All eventually positive solutions of (1.1) *are* q*-rapidly varying provided* (3.3) *holds.*

(iv) *If* (1.1) *is nonoscillatory (which can be guaranteed, e.g., by* $t^\alpha p(t) \le \omega_q/q^{\alpha-1}$ *for large* t*) and*

$$\liminf_{t\to\infty} t^\alpha p(t) > -\infty, \tag{3.4}$$

then all eventually positive solutions of (1.1) *are* q*-regularly bounded.*

Conversely, if there exists an eventually positive solution u *of* (1.1) *such that* $u \in \mathcal{RB}_q$, *then*

$$-\infty < \liminf_{t\to\infty} t^\alpha p(t) \le \limsup_{t\to\infty} t^\alpha p(t) < \frac{1+q^{1-\alpha}}{(q-1)^\alpha}. \tag{3.5}$$

If, in addition, p *is eventually positive or* u *is eventually increasing, then the constant on the right-hand side of* (3.5) *can be improved to* $1/(q-1)^\alpha$.

Proof. (i) *Necessity*. Assume that u is a solution of (1.1) such that $u \in \mathcal{RV}_q(\vartheta_1)$. Then, by Lemma 2.1,

$$\begin{aligned}
\lim_{t\to\infty} t^\alpha p(t) &= -(q-1)^{-\alpha} \lim_{t\to\infty} \mathcal{L}[u](t) = -(q-1)^{-\alpha} \lim_{t\to\infty} F\left(\frac{u(qt)}{u(t)}\right) \\
&= -(q-1)^{-\alpha} F\left(q^{\vartheta_1}\right) = -[1-\alpha]_q \left[\Phi\left([\vartheta_1]_q\right) - h\left(\Phi[\vartheta_1]_q\right)\right] \\
&= \frac{[1-\alpha]_q P}{[1-\alpha]_q} = P.
\end{aligned} \tag{3.6}$$

The same arguments work when dealing with $v \in \mathcal{RV}_q(\vartheta_2)$ instead of u.

to Contents, page 1

Sufficiency. Assume that (3.1) holds. Then there exist $N \in [0,\infty)$, $t_0 \in q^{\mathbb{N}_0}$, and $P_\eta \in (0,\omega_q/q^{\alpha-1})$ such that $-N \le t^\alpha p(t) \le P_\eta$ for $t \in [t_0,\infty)_q$. Let $\mathcal{X}$ be the Banach space of all bounded functions $[t_0,\infty)_q \to \mathbb{R}$ endowed with the supremum norm. Denote $\Omega = \{w \in \mathcal{X} : \Phi(q^{-\eta}-1) \le w(t) \le \widetilde{N}$ for $t \in [t_0,\infty)_q\}$, where $\widetilde{N} = N(q-1)^\alpha + q^{1-\alpha}$, $\eta = \log_q[(q-1)\Phi^{-1}(\lambda_\eta)+1]$, λ_η being the smaller root of $\lambda = h(\lambda) - P_\eta/[1-\alpha]_q$. In view of Lemma 2.1, it holds that $\eta < (\alpha-1)/\alpha$. Moreover, if $P_\eta \ge P$ (which must be valid in our case), then $\vartheta_1 \le \eta$. Further, by Lemma 2.1, $-(q-1)_\eta^P = \Phi(q^{-\eta}-1)(1-q^{(\alpha-1)(\eta-1)})$. Let $\mathcal{T} : \Omega \to \mathcal{X}$ be the operator defined by

$$(\mathcal{T}w)(t) = -(q-1)^\alpha t^\alpha p(t) - \Phi\left(\frac{1}{q\Phi^{-1}(w(qt)) + q} - \frac{1}{q}\right). \tag{3.7}$$

By means of the contraction mapping theorem we will prove that $\mathcal{T}$ has a fixed-point in Ω. First we show that $\mathcal{T}\Omega \subseteq \Omega$. Let $w \in \Omega$. Then, using identities (v) and (vi) from Lemma 2.1,

$$\begin{aligned}(\mathcal{T}w)(t) &\ge -(q-1)^\alpha P_\eta - \Phi\left(\frac{1}{qq^{-\eta}} - \frac{1}{q}\right)\\ &= (\lambda_\eta - h(\lambda_\eta))(q-1)^\alpha[1-\alpha]_q - q^{(\eta-1)(\alpha-1)}\Phi(1-q^{-\eta})\\ &= F(q^\eta) - q^{(\eta-1)(\alpha-1)}\Phi(1-q^{-\eta})\\ &= \Phi(q^{-\eta}-1)\left(1-q^{(\alpha-1)(\eta-1)}\right) - q^{(\eta-1)(\alpha-1)}\Phi(1-q^{-\eta})\\ &= \Phi(q^{-\eta}-1)\end{aligned} \tag{3.8}$$

and $(\mathcal{T}w)(t) \le -(q-1)^\alpha t^\alpha p(t) + q^{1-\alpha} \le \widetilde{N}$ for $t \in [t_0,\infty)_q$. Now we prove that $\mathcal{T}$ is a contraction mapping on Ω. Consider the function $g : (-1,\infty) \to \mathbb{R}$ defined by $g(x) = -\Phi(1/(q\Phi^{-1}(x) + q) - 1/q)$. It is easy to see that $|g'(x)| = q^{1-\alpha}(\Phi^{-1}(x)+1)^{-\alpha}$. Let $w, z \in \Omega$. The Lagrange mean value theorem yields $|g(w(t)) - g(z(t))| = |w(t) - z(t)\|g'(\xi(t))|$, where $\xi : q^{\mathbb{N}_0} \to \mathbb{R}$ is such that $\min\{w(t), z(t)\} \le \xi(t) \le \max\{w(t), z(t)\}$ for $t \in [t_0,\infty)_q$. Hence,

$$\begin{aligned}|(\mathcal{T}w)(t) - (\mathcal{T}z)(t)| &= |g(w(qt)) - g(z(qt))|\\ &= |w(qt) - z(qt)||g'(\xi(t))|\\ &\le |w(qt) - z(qt)||g'(\Phi(q^{-\eta}-1))|\\ &= q^{\eta\alpha+1-\alpha}|w(qt) - z(qt)|\\ &\le q^{\eta\alpha+1-\alpha}\|w - z\|\end{aligned} \tag{3.9}$$

for $t \in [t_0,\infty)_q$. Thus $\|\mathcal{T}w - \mathcal{T}z\| \le q^{\eta\alpha+1-\alpha}\|w-z\|$, where $q^{\eta\alpha+1-\alpha} \in (0,1)$ by virtue of $q > 1$ and $\eta < (\alpha-1)/\alpha$. The Banach fixed-point theorem now guarantees the existence of $w \in \Omega$ such that $w = \mathcal{T}w$. Define u by $u(t) = \prod_{s\in[t_0,t)_q}(\Phi^{-1}(w(s)) + 1)^{-1}$. Then u is a positive solution of $\mathcal{L}[u](t) = -(q-1)^\alpha t^\alpha p(t)$ on $[t_0,\infty)_q$, and, consequently, of (1.1) (this implies nonoscillation of (1.1)). Moreover, $q^{-\eta} \le \Phi^{-1}(w(t)) + 1 \le 1/\overline{N}$, where $\overline{N} = 1/(\Phi^{-1}(\widetilde{N}) + 1)$, and thus

$\overline{N} \leq u(qt)/u(t) \leq q^{\eta}$. Denote $M_* = \liminf_{t\to\infty} u(qt)/u(t)$ and $M^* = \limsup_{t\to\infty} u(qt)/u(t)$. Rewrite $\mathcal{L}[u](t) = -(q-1)^{\alpha} t^{\alpha} p(t)$ as

$$\Phi\left(\frac{u(q^2t)}{qu(qt)} - \frac{1}{q}\right) = \Phi\left(1 - \frac{u(t)}{u(qt)}\right) - (q-1)^{\alpha} t^{\alpha} p(t). \tag{3.10}$$

Taking lim inf and lim sup as $t \to \infty$ in (3.10), we get $\Phi(M_*/q - 1/q) = \Phi(1-1/M_*) - (q-1)^{\alpha} P$ and $\Phi(M^*/q - 1/q) = \Phi(1 - 1/M^*) - (q-1)^{\alpha} P$, respectively. Hence, $F(M_*) = F(M^*)$. Since $M_*, M^* \in [\overline{N}, q^{\eta}]$ and F is strictly decreasing on $(0, q^{(\alpha-1)/\alpha})$ (by Lemma 2.1), we have $M := M_* = M^*$. Moreover,

$$F(M) = -(q-1)^{\alpha} P = (q-1)^{\alpha} [1-\alpha]_q \left(\Phi\left([\vartheta_i]_q\right) - h\left(\Phi[\vartheta_i]_q\right)\right) = F\left(q^{\vartheta_i}\right), \tag{3.11}$$

$i = 1, 2$, which implies $M = q^{\vartheta_1}$, in view of the facts that $M, q^{\vartheta_1} \in (0, q^{(\alpha-1)/\alpha})$, $q^{\vartheta_2} > q^{(\alpha-1)/\alpha}$, and F is monotone on $(0, q^{(\alpha-1)/\alpha})$. Thus $u \in \mathcal{RV}_q(\vartheta_1)$. Now we show that there exists a solution v of (1.1) with $v \in \mathcal{RV}_q(\vartheta_2)$. We can assume that N, t_0, and P_η are the same as in the previous part. Consider the set $\Gamma = \{w \in \mathcal{X} : \Phi(q^{\zeta-1} - 1/q) \leq w(t) \leq \widetilde{M}$ for $t \in [t_0, \infty)_q\}$, where $\widetilde{M} = 1 + (q-1)^{\alpha} N$, $\zeta = \log_q[(q-1)\Phi^{-1}(\lambda_\zeta) + 1]$, λ_ζ being the larger root of $\lambda = h(\lambda) - P_\eta/[1-\alpha]_q$. It is clear that N can be chosen in such a way that $\Phi(q^{\vartheta_2 - 1} - 1/q) < \widetilde{M}$. It holds $(\alpha-1)/\alpha < \zeta \leq \vartheta_2$ and $-(q-1)^{\alpha} P_\eta = \Phi(q^{-\zeta} - 1)(1 - q^{(\alpha-1)(\zeta-1)})$. Define $\mathcal{S} : \Gamma \to \mathcal{X}$ by $(\mathcal{S}w)(t) = \Phi(1 - 1/(q\Phi^{-1}(w(t/q)) + 1)) - (q-1)^{\alpha} t^{\alpha} p(t)$ for $t \in [qt_0, \infty)_q$, and $(\mathcal{S}w)(t_0) = \Phi(q^{\vartheta_2 - 1} - 1/q)$. Using similar arguments as above it is not difficult to see that $\mathcal{S}\Gamma \subseteq \Gamma$ and $\|\mathcal{S}w - \mathcal{S}z\| < q^{\alpha-1-\alpha\zeta} \|w - z\|$ for $w, z \in \Gamma$. So there exists $w \in \Gamma$ such that $w = \mathcal{S}w$. If we define $v(t) = \prod_{s \in [qt_0, t)_q} (q\Phi^{-1}(w(s/q)) + 1)$, then v is a positive solution of (1.1) on $[qt_0, \infty)_q$, which satisfies $q^{\zeta} \leq v(qt)/v(t) \leq q\Phi^{-1}(\widetilde{M}) + 1$. Arguing as above we show that $v \in \mathcal{RV}_q(\vartheta_2)$.

(ii) *Necessity*. The proof is similar to that of (i).

Sufficiency. The condition $t^{\alpha} p(t) \leq \omega_q / q^{\alpha-1}$ for large t implies nonoscillation of (1.1). Indeed, it is easy to see that $y(t) = t^{(\alpha-1)/\alpha}$ is a nonoscillatory solution of the Euler type equation $D_q(\Phi(D_q y(t))) + \omega_q q^{1-\alpha} t^{-\alpha} \Phi(y(qt)) = 0$. Nonoscillation of (1.1) then follows by using the Sturm type comparison theorem, see also Section 4(i). Let us write P as $P = [1-\alpha]_q (h(\Phi([(\alpha-1)/\alpha]_q)) - \Phi([(\alpha-1)/\alpha]_q))$, with noting that $\lambda = \Phi([(\alpha-1)/\alpha]_q)$ is the double root of $\lambda = h(\lambda) - \omega_q q^{1-\alpha}/[1-\alpha]_q$, see Lemma 2.1. Then, in view of Lemma 2.1, we obtain

$$\begin{aligned} F\left(q^{\vartheta}\right) &= (q-1)^{\alpha} [1-\alpha]_q \left[\Phi\left([\vartheta]_q\right) - h\left(\Phi[\vartheta]_q\right)\right] = -\frac{(q-1)^{\alpha} \omega_q}{q^{\alpha-1}} \\ &= -(q-1)^{\alpha} \lim_{t\to\infty} t^{\alpha} p(t) = \lim_{t\to\infty} \mathcal{L}[u](t). \end{aligned} \tag{3.12}$$

Let us denote $U_* = \liminf_{t\to\infty} u(qt)/u(t)$ and $U^* = \limsup_{t\to\infty} u(qt)/u(t)$. It is impossible to have $U_* = 0$ or $U^* = \infty$, otherwise $\lim_{t\to\infty} \mathcal{L}[u](t) = \infty$, which contradicts to (3.12). Thus $0 < U_* \leq U^* < \infty$. Consider (1.1) in the form (3.10). Taking lim sup, respectively, lim inf as $t \to \infty$ in (3.10), into which our u is plugged, we obtain $F(U_*) = F(q^{(\alpha-1)/\alpha}) = F(U^*)$. Thanks to the properties of F, see Lemma 2.1, we get $U_* = U^* = q^{(\alpha-1)/\alpha}$. Hence, $u \in \mathcal{RV}_q((\alpha-1)/\alpha)$.

to Contents, page 1

Since we worked with an arbitrary positive solution, it implies that all positive solutions must be q-regularly varying of index $(\alpha-1)/\alpha$.

(iii) The proof repeats the same arguments as that of [3, Theorem 1] (in spite of no sign condition on p). Note just that condition (3.3) compels p to be eventually negative and the proof of necessity does not depend on the sign of p.

(iv) *Sufficiency*. Let u be an eventually positive solution of (1.1). Assume by a contradiction that $\limsup_{t\to\infty} y(qt)/y(t)=\infty$. Then, in view of Lemma 2.1(vii),

$$\infty=\limsup_{t\to\infty}\left(\Phi\left(\frac{y(q^2t)}{qy(qt)}-\frac{1}{q}\right)-1\right)\leq\limsup_{t\to\infty}\mathcal{L}[y](t)=-(q-1)^{\alpha}\liminf_{t\to\infty}t^{\alpha}p(t)<\infty \quad (3.13)$$

by (3.4), a contradiction. If $\liminf_{t\to\infty} y(qt)/y(t)=0$, then $\limsup_{t\to\infty} y(t)/y(qt)=\infty$ and we proceed similarly as in the previous case. Since we worked with an arbitrary positive solution, it implies that all positive solutions must be q-regularly bounded.

Necessity. Let $y\in\mathcal{RB}_q$ be a solution of (1.1). Taking lim sup as $t\to\infty$ in $-(q-1)^{\alpha}t^{\alpha}p(t)=\mathcal{L}[y](t)$, we get

$$\begin{aligned}&-(q-1)^{\alpha}\liminf_{t\to\infty}t^{\alpha}p(t)\\&\quad=\limsup_{t\to\infty}\mathcal{L}[y](t)\leq\limsup_{t\to\infty}\Phi\left(\frac{y(q^2t)}{qy(qt)}-\frac{1}{q}\right)+\limsup_{t\to\infty}\Phi\left(\frac{y(t)}{y(qt)}-1\right)<\infty,\end{aligned}\quad (3.14)$$

which implies the first inequality in (3.5). Similarly, the lim inf as $t\to\infty$ yields $-(q-1)^{\alpha}\limsup_{t\to\infty}t^{\alpha}p(t)>-1/q^{\alpha-1}-1$, which implies the last inequality in (3.5). If p is eventually positive, then every eventually positive solution of (1.1) is eventually increasing, which can be easily seen from its concavity. Hence, $y(qt)/y(t)\geq 1$ for large t. Thus the last inequality becomes $-(q-1)^{\alpha}\limsup_{t\to\infty}t^{\alpha}p(t)>-1$. □

We are ready to provide a summarizing thorough discussion on asymptotic behavior of solutions to (1.1) with respect to the limit behavior of $t^{\alpha}p(t)$ in the framework of q-Karamata theory. Denote

$$P=\lim_{t\to\infty}t^{\alpha}p(t),\qquad P_*=\liminf_{t\to\infty}t^{\alpha}p(t),\qquad P^*=\limsup_{t\to\infty}t^{\alpha}p(t). \quad (3.15)$$

The set of all q-regularly varying and q-rapidly varying functions is said to be *q-Karamata functions*. With the use of the previous results we obtain the following statement.

Corollary 3.2. (i) *Assume that there exists* $P\in\mathbb{R}\cup\{-\infty,\infty\}$. *In this case,* (1.1) *possesses solutions that are q-Karamata functions provided* (1.1) *is nonoscillatory. Moreover, we distinguish the following subcases:*

(a) $P=-\infty$*:* (1.1) *is nonoscillatory and all its positive solutions are q-rapidly varying (of index* $-\infty$ *or* ∞*).*

(b) $P\in(-\infty,\omega_q/q^{\alpha-1})$*:* (1.1) *is nonoscillatory and there exist a positive solution which is q-regularly varying of index* ϑ_1 *and a positive solution which is q-regularly varying of index* ϑ_2.

to Contents, page 1

(c) $P = \gamma_q$: (1.1) *either oscillatory or nonoscillatory (the latter one can be guaranteed, e.g., by* $t^\alpha p(t) \le \omega_q/q^{\alpha-1}$ *for large* t*). In case of nonoscillation of* (1.1) *all its positive solutions are* q*-regularly varying of index* $(\alpha-1)/\alpha$.

(d) $P \in (\omega_q/q^{\alpha-1}, \infty) \cup \{\infty\}$: (1.1) *is oscillatory.*

(ii) *Assume that* $\mathbb{R} \cup \{-\infty\} \ni P_* < P^* \in \mathbb{R} \cup \{\infty\}$. *In this case, there are no* q*-Karamata functions among positive solutions of* (1.1). *Moreover, we distinguish the following subcases:*

(a) $P_* \in (\omega_q/q^{\alpha-1}, \infty) \cup \{\infty\}$: (1.1) *is oscillatory.*

(b) $P_* \in \{-\infty\} \cup (-\infty, \omega_q/q^{\alpha-1}]$: (1.1) *is either oscillatory (this can be guaranteed, e.g., by* $P^* > (1+q^{1-\alpha})/(q-1)^\alpha$ *or by* $p > 0$ *and* $P^* \ge 1/(q-1)^\alpha$*) or nonoscillatory (this can be guaranteed, e.g., by* $t^\alpha p(t) \le \omega_q/q^{\alpha-1}$ *for large* t*). If, in addition to nonoscillation of* (1.1)*, it holds* $P_* > -\infty$*, then all its positive solutions are* q*-regularly bounded, but there is no* q*-regularly varying solution. If* $P_* = -\infty$*, then there is no* q*-regularly bounded or* q*-rapidly varying solution.*

4. Concluding Remarks

(i) We start with some remarks to Kneser type criteria. As a by product of Theorem 3.1(i) we get the following nonoscillation Kneser type criterion: if $\lim_{t\to\infty} t^\alpha p(t) < \omega_q/q^{\alpha-1}$, then (1.1) is nonoscillatory. However, its better variant is known (it follows from a more general time-scale case involving Hille-Nehari type criterion [15]), where the sufficient condition is relaxed to $\limsup_{t\to\infty} t^\alpha p(t) < \omega_q/q^{\alpha-1}$. The constant $\omega_q/q^{\alpha-1}$ is sharp, since $\liminf_{t\to\infty} t^\alpha p(t) > \omega_q/q^{\alpha-1}$ implies oscillation of (1.1), see [15]. But no conclusion can be generally drawn if the equality occurs in these conditions. The above lim sup nonoscillation criterion can be alternatively obtained also from the observation presented at the beginning of the proof of Theorem 3.1(ii) involving the Euler type q-difference equation. And it is worthy of note that the conclusion of that observation can be reached also when modifying the proof of Hille-Nehari type criterion in [15]. A closer examination of the proof of Theorem 3.1(iv) shows that a necessary condition for nonoscillation of (1.1) is $-(q-1)^\alpha \limsup_{t\to\infty} t^\alpha p(t) \ge -q^{1-\alpha} - 1$. Thus we have obtained quite new Kneser type oscillation criterion: if $\limsup_{t\to\infty} t^\alpha p(t) > (1+q^{1-\alpha})/(q-1)^\alpha$, then (1.1) is oscillatory. If p is eventually positive, then the constant on the right-hand side can be improved to $1/(q-1)^\alpha$ and the strict inequality can be replaced by the nonstrict one (this is because of q-regular boundedness of possible positive solutions). A continuous analog of this criterion is not known, which is quite natural since $1/(q-1)^\alpha \to \infty$ as $q \to 1$. Compare these results with the Hille-Nehari type criterion, which was proved in general setting for dynamic equations and time-scales, and is valid no matter what the graininess is (see [15]); in q-calculus it reads as follows: if $p \ge 0$ and $\limsup_{t\to\infty} t^{\alpha-1}\int_t^\infty p(s)\mathrm{d}_q s > 1$, then (1.1) is oscillatory. This criterion holds literally also in the continuous case. Finally note that, in general, $\limsup_{t\to\infty} t^{\alpha-1}\int_t^\infty p(s)\mathrm{d}_q s \le \limsup_{t\to\infty} -[1-\alpha]_q t^\alpha p(t)$.

(ii) The results contained in Theorem 3.1 can understood at least in the three following ways:

(a) As a q-version of the continuous results for (1.2) from [5]. However, there are several substantial differences: The conditions in the continuous case are (and somehow must be) in the integral form (see also the item (iii) of this section); there is a different approach in the proof (see also the item (iv) of this section); the rapid variation has not been treated in such detail in the continuous case; in the case of the

to Contents, page 1

existence of the double root, we show that all (and not just some) positive solutions are q-regularly varying under quite mild assumptions; for positive solutions to be q-regularly bounded we obtain quite simple and natural sufficient and also necessary conditions.

(b) As a half-linear extension of the results for $D_q^2 y(t) + p(t)y(qt) = 0$ from [1]. In contrast to the linear case, in the half-linear case a reduction of order formula is not at disposal. Thus to prove that there are two q-regularly varying solutions of two different indices we need immediately to construct both of them. Lack of a fundamental like system for half-linear equations causes that, for the time being, we are not able to show that all positive solutions are q-regularly varying. This is however much easier task when $p(t) < 0$, see [3].

(c) As a generalization of the results from [3] in the sense of no sign condition on the coefficient p.

(iii) From the continuous theory we know that the sufficient and necessary conditions for regularly or rapidly varying behavior of solutions to (1.2) are in terms of limit behavior of integral expressions, typically $t^{\alpha-1}\int_t^{\infty} p(s)\mathrm{d}s$ or $t^{\alpha-1}\int_t^{\lambda t} p(s)\mathrm{d}s$. In contrast to that, in q-calculus case the conditions have nonintegral form. This is the consequence of specific properties of q-calculus: one thing is that we use a different approach which does not apply in the continuous case. Another thing is that the limit $\lim_{t\to\infty} t^{\alpha-1}\int_t^{\infty} p(s)\mathrm{d}_q s$ can be expressed in terms of $\lim_{t\to\infty} t^{\alpha} p(t)$ (and vice versa), provided it exists. Such a relation does not work in the continuous case.

(iv) As already said, our approach in the proof of Theorem 3.1 is different from what is known in the continuous theory. Our method is designed just for q-difference equations and roughly speaking, it is based on rewriting a q-difference equation in terms of the fractions which appear in Definition 2.2. Such a technique cannot work in the continuous case. Since this method uses quite natural and simple relations (which are possible thanks to the special structure of $q^{\mathbb{N}_0}$), we believe that it will enable us to prove also another results which are q-versions of existing or nonexisting continuous results; in the latter case, such results may serve to predict a possible form of the continuous counterpart, which may be difficult to handle directly. We just take, formally, the limit as $q \to 1+$.

Acknowledgments

The author was supported by Grant 201/10/1032 of the Czech Grant Agency and by the Institutional Research plan AV0Z010190503.

References

[1] P. Řehák, "Second order linear q-difference equations: nonoscillation and asymptotics," submitted.

[2] P. Řehák and J. Vítovec, "q-regular variation and q-difference equations," *Journal of Physics A*, vol. 41, no. 49, Article ID 495203, 10 pages, 2008.

[3] P. Řehák and J. Vítovec, "q-Karamata functions and second order q-difference equations," submitted.

[4] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular Variation*, vol. 27 of *Encyclopedia of Mathematics and Its Applications*, Cambridge University Press, Cambridge, 1987.

[5] J. Jaroa, T. Kusano, and T. Tanigawa, "Nonoscillation theory for second order half-linear differential equations in the framework of regular variation," *Results in Mathematics*, vol. 43, no. 1-2, pp. 129–149, 2003.

to Contents, page 1

[6] P. Řehák and J. Vítovec, "Regular variation on measure chains," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 72, no. 1, pp. 439–448, 2010.
[7] O. Došlý and P. Řehák, *Half-Linear Differential Equations*, vol. 202 of *North-Holland Mathematics Studies*, Elsevier Science B.V., Amsterdam, The Netherlands, 2005.
[8] V. Marić, *Regular Variation and Differential Equations*, vol. 1726 of *Lecture Notes in Mathematics*, Springer, Berlin, Germany, 2000.
[9] G. Bangerezako, "An introduction to q-difference equations," preprint.
[10] G. D. Birkhoff and P. E. Guenther, "Note on a canonical form for the linear q-difference system," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 27, pp. 218–222, 1941.
[11] M. Bohner and M. Ünal, "Kneser's theorem in q-calculus," *Journal of Physics A*, vol. 38, no. 30, pp. 6729–6739, 2005.
[12] V. Kac and P. Cheung, *Quantum Calculus*, Universitext, Springer, New York, NY, USA, 2002.
[13] G. Gasper and M. Rahman, *Basic Hypergeometric Series*, vol. 96 of *Encyclopedia of Mathematics and Its Applications*, Cambridge University Press, Cambridge, UK, 2nd edition, 2004.
[14] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales: An Introduction with Applications*, Birkhäuser, Boston, Mass, USA, 2001.
[15] P. Řehák, "A critical oscillation constant as a variable of time scales for half-linear dynamic equations," *Mathematica Slovaca*, vol. 60, no. 2, pp. 237–256, 2010.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 981401, 9 pages
doi:10.1155/2011/981401

Research Article

Asymptotic Formula for Oscillatory Solutions of Some Singular Nonlinear Differential Equation

Irena Rachůnková and Lukáš Rachůnek

Department of Mathematics, Faculty of Science, Palacký University, 17. Listopadu 12, 771 46 Olomouc, Czech Republic

Correspondence should be addressed to Irena Rachůnková, irena.rachunkova@upol.cz

Received 28 October 2010; Revised 31 March 2011; Accepted 2 May 2011

Academic Editor: Yuri V. Rogovchenko

Singular differential equation $(p(t)u')' = p(t)f(u)$ is investigated. Here f is Lipschitz continuous on $\mathbb{R}$ and has at least two zeros 0 and $L > 0$. The function p is continuous on $[0,\infty)$ and has a positive continuous derivative on $(0,\infty)$ and $p(0) = 0$. An asymptotic formula for oscillatory solutions is derived.

1. Introduction

In this paper, we investigate the equation

$$\left(p(t)u'\right)' = p(t)f(u), \quad t \in (0,\infty), \tag{1.1}$$

where f satisfies

$$f \in Lip_{\mathrm{loc}}(\mathbb{R}), \quad f(0) = f(L) = 0, \; f(x) < 0, \; x \in (0,L), \tag{1.2}$$

$$\exists \overline{B} \in (-\infty,0)\colon f(x) > 0, \quad x \in \left[\overline{B},0\right), \tag{1.3}$$

$$F\left(\overline{B}\right) = F(L), \quad \text{where } F(x) = -\int_0^x f(z)\mathrm{d}z, \; x \in \mathbb{R}, \tag{1.4}$$

and p fulfils

$$p \in C[0,\infty) \cap C^1(0,\infty), \quad p(0) = 0, \tag{1.5}$$

$$p'(t) > 0, \quad t \in (0,\infty), \quad \lim_{t\to\infty}\frac{p'(t)}{p(t)} = 0. \tag{1.6}$$

to Contents, page 1

Equation (1.1) is a generalization of the equation

$$u'' + \frac{k-1}{t}u' = f(u), \quad t \in (0,\infty), \tag{1.7}$$

which arises for $k > 1$ and special forms of f in many areas, for example: in the study of phase transitions of Van der Waals fluids [1–3], in population genetics, where it serves as a model for the spatial distribution of the genetic composition of a population [4, 5], in the homogeneous nucleation theory [6], in the relativistic cosmology for the description of particles which can be treated as domains in the universe [7], in the nonlinear field theory, in particular, when describing bubbles generated by scalar fields of the Higgs type in the Minkowski spaces [8]. Numerical simulations of solutions of (1.1), where f is a polynomial with three zeros have been presented in [9–11]. Close problems about the existence of positive solutions can be found in [12–14].

Due to $p(0) = 0$, (1.1) has a singularity at $t = 0$.

Definition 1.1. A function $u \in C^1[0,\infty) \cap C^2(0,\infty)$ which satisfies (1.1) for all $t \in (0,\infty)$ is called a *solution* of (1.1).

Definition 1.2. Let u be a solution of (1.1) and let L be of (1.2). Denote $u_{\text{sup}} = \sup\{u(t) : t \in [0,\infty)\}$. If $u_{\text{sup}} < L$ ($u_{\text{sup}} = L$ or $u_{\text{sup}} > L$), then u is called a *damped* solution (a *bounding homoclinic* solution or an *escape* solution).

These three types of solutions have been investigated in [15–19]. In particular, the existence of damped oscillatory solutions which converge to 0 has been proved in [19].

The main result of this paper is contained in Section 3 in Theorem 3.1, where we provide an asymptotic formula for damped oscillatory solutions of (1.1).

2. Existence of Oscillatory Solutions

Here, we will study solutions of (1.1) satisfying the initial conditions

$$u(0) = B, \qquad u'(0) = 0, \tag{2.1}$$

with a parameter $B \leq L$. Reason is that we focus our attention on damped solutions of (1.1) and that each solution u of (1.1) must fulfil $u'(0) = 0$ (see [19]).

First, we bring two theorems about the existence of damped and oscillatory solutions.

Theorem 2.1 (see [19]). *Assume that* (1.2)–(1.6) *hold. Then for each* $B \in [\overline{B}, L)$ *problem* (1.1), (2.1) *has a unique solution. This solution is damped.*

Theorem 2.2. *Assume that* (1.2)–(1.6) *hold. Further, let there exists* $k_0 \in (0,\infty)$ *such that*

$$p \in C^2(0,\infty), \qquad \limsup_{t\to\infty}\left|\frac{p''(t)}{p'(t)}\right| < \infty, \qquad \liminf_{t\to\infty}\frac{p(t)}{t^{k_0}} \in (0,\infty], \tag{2.2}$$

$$\lim_{x\to 0+}\frac{f(x)}{x} < 0, \qquad \lim_{x\to 0-}\frac{f(x)}{x} < 0. \tag{2.3}$$

to Contents, page 1

Then for each $B \in [\overline{B}, L)$ *problem* (1.1), (2.1) *has a unique solution* u*. If* $B \neq 0$*, then the solution* u *is damped and oscillatory with decreasing amplitudes and*

$$\lim_{t\to\infty} u(t) = 0. \tag{2.4}$$

Proof. The assertion follows from Theorems 2.3, 2.10 and 3.1 in [19]. □

Example 2.3. The functions

(i) $p(t) = t^k$, $p(t) = t^k \ln(t^\ell + 1)$, $k, \ell \in (0, \infty)$,

(ii) $p(t) = t + \alpha \sin t$, $\alpha \in (-1, 1)$,

(iii) $p(t) = t^k/(1 + t^\ell)$, $k, \ell \in (0, \infty)$, $\ell < k$

satisfy (1.5), (1.6), and (2.2).

The functions

(i) $p(t) = \ln(t + 1)$, $p(t) = \arctan t$, $p(t) = t^k/(1 + t^k)$, $k \in (0, \infty)$

satisfy (1.5), (1.6), but not (2.2) (the third condition).

The function

(i) $p(t) = t^k + \alpha \sin t^k$, $\alpha \in (-1, 1)$, $k \in (1, \infty)$,

satisfy (1.5), (1.6) but not (2.2) (the second and third conditions).

Example 2.4. Let $k \in (0, \infty)$.

(i) The function

$$f(x) = \begin{cases} -kx, & \text{for } x \le 0, \\ x(x-1), & \text{for } x > 0, \end{cases} \tag{2.5}$$

satisfies (1.2) with $L = 1$, (1.3), (1.4) with $\overline{B} = -(3k)^{-1/2}$ and (2.3).

(ii) The function

$$f(x) = \begin{cases} kx^2, & \text{for } x \le 0, \\ x(x-1), & \text{for } x > 0, \end{cases} \tag{2.6}$$

satisfies (1.2) with $L = 1$, (1.3), (1.4) with $\overline{B} = -(2k)^{-1/3}$ but not (2.3) (the second condition).

In the next section, the generalized Matell's theorem which can be found as Theorem 6.5 in the monograph by Kiguradze will be useful. For our purpose, we provide its following special case.

Consider an interval $J \subset \mathbb{R}$. We write $AC(J)$ for the set of functions absolutely continuous on J and $AC_{\text{loc}}(J)$ for the set of functions belonging to $AC(I)$ for each compact

to Contents, page 1

interval $I \subset J$. Choose $t_0 > 0$ and a function matrix $A(t) = (a_{i,j}(t))_{i,j\le 2}$ which is defined on (t_0, ∞). Denote by $\lambda(t)$ and $\mu(t)$ eigenvalues of $A(t)$, $t \in (t_0, \infty)$. Further, suppose

$$\lambda = \lim_{t\to\infty} \lambda(t), \qquad \mu = \lim_{t\to\infty} \mu(t) \tag{2.7}$$

be different eigenvalues of the matrix $A = \lim_{t\to\infty} A(t)$, and let $\mathbf{l}$ and $\mathbf{m}$ be eigenvectors of A corresponding to λ and μ, respectively.

Theorem 2.5 (see [20]). *Assume that*

$$a_{i,j} \in AC_{\text{loc}}(t_0, \infty), \qquad \left| \int_{t_0}^{\infty} a'_{i,j}(t)dt \right| < \infty, \ i, j = 1, 2, \tag{2.8}$$

and that there exists $c_0 > 0$ such that

$$\int_s^t \operatorname{Re}(\lambda(\tau) - \mu(\tau))\mathrm{d}\tau \le c_0, \quad t_0 \le s < t, \tag{2.9}$$

or

$$\int_{t_0}^{\infty} \operatorname{Re}(\lambda(\tau) - \mu(\tau))\mathrm{d}\tau = \infty, \qquad \int_s^t \operatorname{Re}(\lambda(\tau) - \mu(\tau))\mathrm{d}\tau \ge -c_0, \quad t_0 \le s < t. \tag{2.10}$$

Then the differential system

$$\mathbf{x}'(t) = A(t)\mathbf{x}(t) \tag{2.11}$$

has a fundamental system of solutions $x(t)$, $y(t)$ such that

$$\lim_{t\to\infty} \mathbf{x}(t) e^{-\int_{t_0}^t \lambda(\tau)\mathrm{d}\tau} = \mathbf{l}, \qquad \lim_{t\to\infty} \mathbf{y}(t) e^{-\int_{t_0}^t \mu(\tau)\mathrm{d}\tau} = \mathbf{m}. \tag{2.12}$$

3. Asymptotic Formula

In order to derive an asymptotic formula for a damped oscillatory solution u of problem (1.1), (2.1), we need a little stronger assumption than (2.3). In particular, the function $f(x)/x$ should have a negative derivative at $x = 0$.

Theorem 3.1. *Assume that* (1.2)–(1.6), *and* (2.2) *hold. Assume, moreover, that there exist $\eta > 0$ and $c > 0$ such that*

$$\frac{f(x)}{x} \in AC[-\eta, \eta], \qquad \lim_{x\to 0} \frac{f(x)}{x} = -c. \tag{3.1}$$

to Contents, page 1

Then for each $B \in [\overline{B}, L)$ *problem* (1.1), (2.1) *has a unique solution* u*. If* $B \neq 0$*, then the solution* u *is damped and oscillatory with decreasing amplitudes such that*

$$\limsup_{t\to\infty} \sqrt{p(t)}|u(t)| < \infty. \tag{3.2}$$

Proof. We have the following steps:

Step 1 (construction of an auxiliary linear differential system). Choose $B \in [\overline{B}, L)$, $B \neq 0$. By Theorem 2.2, problem (1.1), (2.1) has a unique oscillatory solution u with decreasing amplitudes and satisfying (2.4). Having this solution u, define a linear differential equation

$$v'' + \frac{p'(t)}{p(t)}v' = \frac{f(u(t))}{u(t)}v, \tag{3.3}$$

and the corresponding linear differential system

$$x_1' = x_2, \qquad x_2' = \frac{f(u(t))}{u(t)}x_1 - \frac{p'(t)}{p(t)}x_2. \tag{3.4}$$

Denote

$$A(t) = \left(a_{i,j}(t)\right)_{i,j\le 2} = \begin{pmatrix} 0 & 1 \\ \dfrac{f(u(t))}{u(t)} & -\dfrac{p'(t)}{p(t)} \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & 1 \\ -c & 0 \end{pmatrix}. \tag{3.5}$$

By (1.6), (2.4), and (3.1),

$$A = \lim_{t\to\infty} A(t). \tag{3.6}$$

Eigenvalues of A are numbers $\lambda = i\sqrt{c}$ and $\mu = -i\sqrt{c}$, and eigenvectors of A are $\mathbf{l} = (1, i\sqrt{c})$ and $\mathbf{m} = (1, -i\sqrt{c})$, respectively. Denote

$$D(t) = \left(\frac{p'(t)}{2p(t)}\right)^2 + \frac{f(u(t))}{u(t)}, \quad t \in (0,\infty). \tag{3.7}$$

Then eigenvalues of $A(t)$ have the form

$$\lambda(t) = -\frac{p'(t)}{2p(t)} + \sqrt{D(t)}, \qquad \mu(t) = -\frac{p'(t)}{2p(t)} - \sqrt{D(t)}, \quad t \in (0,\infty). \tag{3.8}$$

We see that

$$\lim_{t\to\infty} \lambda(t) = \lambda, \qquad \lim_{t\to\infty} \mu(t) = \mu. \tag{3.9}$$

to Contents, page 1

Step 2 (verification of the assumptions of Theorem 2.5). Due to (1.6), (2.4), and (3.1), we can find $t_0 > 0$ such that

$$u(t_0) \neq 0, \qquad |u(t)| \leq \eta, \qquad D(t) < 0, \quad t \in (t_0, \infty). \tag{3.10}$$

Therefore, by (3.1),

$$a_{21}(t) = \frac{f(u(t))}{u(t)} \in AC_{\mathrm{loc}}(t_0, \infty), \tag{3.11}$$

and so

$$\left| \int_{t_0}^{\infty} \left(\frac{f(u(t))}{u(t)} \right)' \mathrm{d}t \right| = \left| \lim_{t \to \infty} \frac{f(u(t))}{u(t)} - \frac{f(u(t_0))}{u(t_0)} \right| = \left| -c - \frac{f(u(t_0))}{u(t_0)} \right| < \infty. \tag{3.12}$$

Further, by (2.2), $a_{22}(t) = -p'(t)/p(t) \in C^1(t_0, \infty)$. Hence, due to (1.6),

$$\left| \int_{t_0}^{\infty} \left(\frac{p'(t)}{p(t)} \right) \mathrm{d}t \right| = \left| \lim_{t \to \infty} \frac{p'(t)}{p(t)} - \frac{p'(t_0)}{p(t_0)} \right| = \frac{p'(t_0)}{p(t_0)} < \infty. \tag{3.13}$$

Since $a_{11}(t) \equiv 0$ and $a_{12}(t) \equiv 1$, we see that (2.8) is satisfied. Using (3.8) we get $\mathrm{Re}(\lambda(t) - \mu(t)) \equiv 0$. This yields

$$\int_s^t \mathrm{Re}(\lambda(\tau) - \mu(\tau)) \mathrm{d}\tau = 0 < c_0, \quad t_0 \leq s < t, \tag{3.14}$$

for any positive constant c_0. Consequently (2.9) is valid.

Step 3 (application of Theorem 2.5). By Theorem 2.5 there exists a fundamental system $\mathbf{x}(t) = (x_1(t), x_2(t))$, $\mathbf{y}(t) = (y_1(t), y_2(t))$ of solutions of (3.4) such that (2.12) is valid. Hence

$$\lim_{t \to \infty} x_1(t) \mathrm{e}^{-\int_{t_0}^t \lambda(\tau) \mathrm{d}\tau} = 1, \qquad \lim_{t \to \infty} y_1(t) \mathrm{e}^{-\int_{t_0}^t \mu(\tau) \mathrm{d}\tau} = 1. \tag{3.15}$$

Using (3.8) and (3.10), we get

$$\begin{aligned} \exp\left(-\int_{t_0}^t \lambda(\tau) \mathrm{d}\tau \right) &= \exp\left(\int_{t_0}^t \left(\frac{p'(\tau)}{2p(\tau)} - \sqrt{D(\tau)} \right) \mathrm{d}\tau \right) \\ &= \exp\left(\frac{1}{2} \ln \frac{p(t)}{p(t_0)} \right) \exp\left(-i \int_{t_0}^t \sqrt{|D(\tau)|} \mathrm{d}\tau \right), \end{aligned} \tag{3.16}$$

to Contents, page 1

and, hence,

$$\left|e^{-\int_{t_0}^{t}\lambda(\tau)d\tau}\right| = \sqrt{\frac{p(t)}{p(t_0)}}, \quad t \in (t_0, \infty). \tag{3.17}$$

Similarly

$$\left|e^{-\int_{t_0}^{t}\mu(\tau)d\tau}\right| = \sqrt{\frac{p(t)}{p(t_0)}}, \quad t \in (t_0, \infty). \tag{3.18}$$

Therefore, (3.15) implies

$$\begin{aligned} 1 &= \lim_{t\to\infty}\left|x_1(t)e^{-\int_{t_0}^{t}\lambda(\tau)d\tau}\right| = \lim_{t\to\infty}|x_1(t)|\sqrt{\frac{p(t)}{p(t_0)}}, \\ 1 &= \lim_{t\to\infty}\left|y_1(t)e^{-\int_{t_0}^{t}\mu(\tau)d\tau}\right| = \lim_{t\to\infty}|y_1(t)|\sqrt{\frac{p(t)}{p(t_0)}}. \end{aligned} \tag{3.19}$$

Step 4 (asymptotic formula). In Step 1, we have assumed that u is a solution of (1.1), which means that

$$u''(t) + \frac{p'(t)}{p(t)}u'(t) = f(u(t)), \quad \text{for } t \in (0, \infty). \tag{3.20}$$

Consequently

$$u''(t) + \frac{p'(t)}{p(t)}u'(t) = \frac{f(u(t))}{u(t)}u(t), \quad \text{for } t \in (0, \infty), \tag{3.21}$$

and, hence, u is also a solution of (3.3). This yields that there are $c_1, c_2 \in \mathbb{R}$ such that $u(t) = c_1x_1(t) + c_2y_1(t),\ t \in (0, \infty)$. Therefore,

$$\limsup_{t\to\infty}\sqrt{p(t)}|u(t)| \le (|c_1| + |c_2|)\sqrt{p(t_0)} < \infty. \tag{3.22}$$

□

Remark 3.2. Due to (2.2) and (3.2), we have for a solution u of Theorem 3.1

$$u(t) = O\left(t^{-k_0/2}\right), \quad \text{for } t \longrightarrow \infty. \tag{3.23}$$

to Contents, page 1

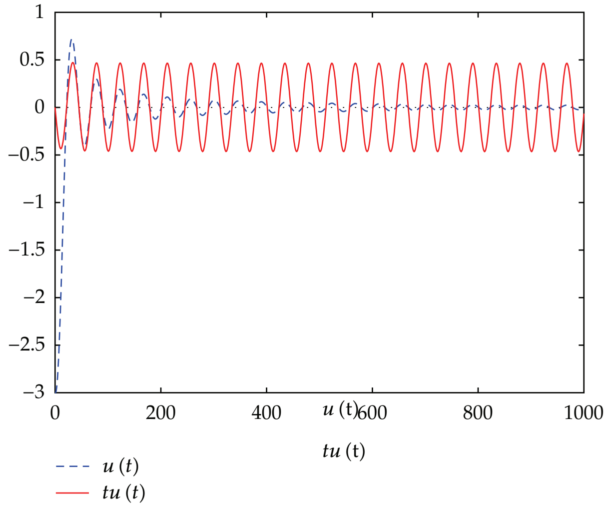


Figure 1

Example 3.3. Let $k \in (1, \infty)$.

(i) The functions $f(x) = x(x-1)$ and $f(x) = x(x-1)(x+2)$ satisfy all assumptions of Theorem 3.1.

(ii) The functions $f(x) = x^{2k-1}(x-1)$ and $f(x) = x^{2k-1}(x-1)(x+2)$

satisfy (1.2)–(1.4) but not (3.1) (the second condition).

Example 3.4. Consider the initial problem

$$\left(t^2 u'\right)' = t^2 u(u-5)(u+10), \qquad u(0) = -3,\ u'(0) = 0. \tag{3.24}$$

Here $L_0 = -10$, $L = 5$ and we can check that $\overline{B} < -3$. Further, all assumptions of Theorems 2.2 and 3.1 are fulfilled. Therefore, by Theorem 2.2, there exists a unique solution u of problem (3.24) which is damped and oscillatory and converges to 0. By Theorem 3.1, we have

$$\limsup_{t\to\infty} t|u(t)| < \infty, \quad \text{that is,}\ \ u(t) = O\left(\frac{1}{t}\right), \quad \text{for } t \longrightarrow \infty. \tag{3.25}$$

The behaviour of the solution $u(t)$ and of the function $tu(t)$ is presented on Figure 1.

Remark 3.5. Our further research of this topic will be focused on a deeper investigation of all types of solutions defined in Definition 1.2. For example, we have proved in [15, 19] that damped solutions of (1.1) can be either oscillatory or they have a finite number of zeros or no zero and converge to 0. A more precise characterization of behaviour of nonoscillatory solutions are including their asymptotic formulas in as open problem. The same can be said about homoclinic solutions. In [17] we have found some conditions which guarantee their existence, and we have shown that if u is a homoclinic solution of (1.1), then $\lim_{t\to\infty} u(t) = L$.

to Contents, page 1

In order to discover other existence conditions for homoclinic solutions, we would like to estimate their convergence by proper asymptotic formulas.

Acknowledgments

The authors thank the referees for comments and suggestions. This paper was supported by the Council of Czech Government MSM 6198959214.

References

[1] V. Bongiorno, L. E. Scriven, and H. T. Davis, “Molecular theory of fluid interfaces,” *Journal of Colloid and Interface Science*, vol. 57, pp. 462–475, 1967.
[2] H. Gouin and G. Rotoli, “An analytical approximation of density profile and surface tension of microscopic bubbles for Van der Waals fluids,” *Mechanics Research Communications*, vol. 24, pp. 255–260, 1997.
[3] J. D. Van Der Waals and R. Kohnstamm, *Lehrbuch der Thermodynamik*, vol. 1, Leipzig, Germany, 1908.
[4] P. C. Fife, *Mathematical Aspects of Reacting and Diffusing Systems*, vol. 28 of *Lecture Notes in Biomathematics*, Springer, Berlin, Germany, 1979.
[5] R. A. Fischer, “The wave of advance of advantegeous genes,” *Journal of Eugenics*, vol. 7, pp. 355–369, 1937.
[6] F. F. Abraham, *Homogeneous Nucleation Theory*, Academies Press, New York, NY, USA, 1974.
[7] A. P. Linde, *Particle Physics and Inflationary Cosmology*, Harwood Academic, Chur, Switzerland, 1990.
[8] G. H. Derrick, “Comments on nonlinear wave equations as models for elementary particles,” *Journal of Mathematical Physics*, vol. 5, pp. 1252–1254, 1964.
[9] F. Dell’Isola, H. Gouin, and G. Rotoli, “Nucleation of spherical shell-like interfaces by second gradient theory: numerical simulations,” *European Journal of Mechanics*, vol. 15, no. 4, pp. 545–568, 1996.
[10] G. Kitzhofer, O. Koch, P. Lima, and E. Weinmüller, “Efficient numerical solution of the density profile equation in hydrodynamics,” *Journal of Scientific Computing*, vol. 32, no. 3, pp. 411–424, 2007.
[11] P. M. Lima, N. V. Chemetov, N. B. Konyukhova, and A. I. Sukov, “Analytical-numerical investigation of bubble-type solutions of nonlinear singular problems,” *Journal of Computational and Applied Mathematics*, vol. 189, no. 1-2, pp. 260–273, 2006.
[12] H. Berestycki, P. L. Lions, and L. A. Peletier, “An ODE approach to the existence of positive solutions for semilinear problems in $\mathbb{R}^N$,” *Indiana University Mathematics Journal*, vol. 30, no. 1, pp. 141–157, 1981.
[13] D. Bonheure, J. M. Gomes, and L. Sanchez, “Positive solutions of a second-order singular ordinary differential equation,” *Nonlinear Analysis: Theory, Methods & Appplications*, vol. 61, no. 8, pp. 1383–1399, 2005.
[14] M. Conti, L. Merizzi, and S. Terracini, “Radial solutions of superlinear equations in $\mathbb{R}^N$, part I: a global variational approach,” *Archive for Rational Mechanics and Analysis*, vol. 153, no. 4, pp. 291–316, 2000.
[15] I. Rachůnková and J. Tomeček, “Bubble-type solutions of nonlinear singular problems,” *Mathematical and Computer Modelling*, vol. 51, no. 5-6, pp. 658–669, 2010.
[16] I. Rachůnková and J. Tomeček, “Strictly increasing solutions of a nonlinear singular differential equation arising in hydrodynamics,” *Nonlinear Analysis: Theory, Methods & Appplications*, vol. 72, no. 3-4, pp. 2114–2118, 2010.
[17] I. Rachůnková and J. Tomeček, “Homoclinic solutions of singular nonautonomous second-order differential equations,” *Boundary Value Problems*, vol. 2009, Article ID 959636, 21 pages, 2009.
[18] I. Rachůnková, J. Tomeček, and J. Stryja, “Oscillatory solutions of singular equations arising in hydrodynamics,” *Advances in Difference Equations*, vol. 2010, Article ID 872160, 13 pages, 2010.
[19] I. Rachůnková, L. Rachůnek, and J. Tomeček, “Existence of oscillatory solutions of singular nonlinear differential equations,” *Abstract and Applied Analysis*, vol. 2011, Article ID 408525, 20 pages, 2011.
[20] I. Kiguradze, *Some Singular Boundary Value Problems for Ordinary Differential Equations*, ITU, Tbilisi, Georgia, 1975.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 730128, 10 pages
doi:10.1155/2011/730128

Research Article

Asymptotic Properties of Third-Order Delay Trinomial Differential Equations

J. Džurina and R. Komariková

Department of Mathematics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Letná 9, 042 00 Košice, Slovakia

Correspondence should be addressed to J. Džurina, jozef.dzurina@tuke.sk

Received 2 September 2010; Accepted 3 November 2010

Academic Editor: Yuri Rogovchenko

The aim of this paper is to study properties of the third-order delay trinomial differential equation $((1/r(t))y''(t))' + p(t)y'(t) + q(t)y(\sigma(t)) = 0$, by transforming this equation onto the second-/third-order binomial differential equation. Using suitable comparison theorems, we establish new results on asymptotic behavior of solutions of the studied equations. Obtained criteria improve and generalize earlier ones.

1. Introduction

In this paper, we will study oscillation and asymptotic behavior of solutions of third-order delay trinomial differential equations of the form

$$\left(\frac{1}{r(t)}\, y''(t)\right)' + p(t)y'(t) + q(t)y(\sigma(t)) = 0. \qquad (E)$$

Throughout the paper, we assume that $r(t), p(t), q(t), \sigma(t) \in C([t_0,\infty))$ and

(i) $r(t) > 0$, $p(t) \geq 0$, $q(t) > 0$, $\sigma(t) > 0$,

(ii) $\sigma(t) \leq t$, $\lim_{t\to\infty}\sigma(t) = \infty$,

(iii) $R(t) = \int_{t_0}^{t} r(s)\, ds \to \infty$ as $t \to \infty$.

By a solution of (E), we mean a function $y(t) \in C^2([T_x,\infty))$, $T_x \geq t_0$, that satisfies (E) on $[T_x,\infty)$. We consider only those solutions $y(t)$ of (E) which satisfy $\sup\{|y(t)| : t \geq T\} > 0$ for all $T \geq T_x$. We assume that (E) possesses such a solution. A solution of (E) is called oscillatory if it has arbitrarily large zeros on $[T_x, \infty)$, and, otherwise, it is nonoscillatory. Equation (E) itself is said to be oscillatory if all its solutions are oscillatory.

Recently, increased attention has been devoted to the oscillatory and asymptotic properties of second- and third-order differential equations (see [1–22]). Various techniques appeared for the investigation of such differential equations. Our method is based on establishing new comparison theorems, so that we reduce the examination of the third-order trinomial differential equations to the problem of the observation of binomial equations.

In earlier papers [11, 13, 16, 20], a particular case of (E), namely, the ordinary differential equation (without delay)

$$y'''(t) + p(t)y'(t) + g(t)y(t) = 0, \tag{E_1}$$

has been investigated, and sufficient conditions for all its nonoscillatory solutions $y(t)$ to satisfy

$$y(t)y'(t) < 0 \tag{1.1}$$

or the stronger condition

$$\lim_{t\to\infty} y(t) = 0 \tag{1.2}$$

are presented. It is known that (E_1) has always a solution satisfying (1.1). Recently, various kinds of sufficient conditions for all nonoscillatory solutions to satisfy (1.1) or (1.2) appeared. We mention here [9, 11, 13, 16, 21]. But there are only few results for differential equations with deviating argument. Some attempts have been made in [8, 10, 18, 19]. In this paper we generalize these, results and we will study conditions under which all nonoscillatory solutions of (E) satisfy (1.1) and (1.2). For our further references we define as following.

Definition 1.1. We say that (E) has property (P_0) if its every nonoscillatory solution $y(t)$ satisfies (1.1).

In this paper, we have two purposes. In the first place, we establish comparison theorems for immediately obtaining results for third-order delay equation from that of third order equation without delay. This part extends and complements earlier papers [7, 8, 10, 18].

Secondly, we present a comparison principle for deducing the desired property of (E) from the oscillation of a second-order differential equation without delay. Here, we generalize results presented in [8, 9, 14, 15, 21].

Remark 1.2. All functional inequalities considered in this paper are assumed to hold eventually;0 that is, they are satisfied for all t large enough.

to Contents, page 1

2. Main Results

It will be derived that properties of (E) are closely connected with the corresponding second-order differential equation

$$\left(\frac{1}{r(t)}\, v'(t)\right)' + p(t)v(t) = 0 \tag{E_v}$$

as the following theorem says.

Theorem 2.1. *Let $v(t)$ be a positive solution of (E_v). Then (E) can be written as*

$$\left(\frac{v^2(t)}{r(t)}\left(\frac{1}{v(t)}\, y'(t)\right)'\right)' + q(t)v(t)y(\sigma(t)) = 0. \tag{E^c}$$

Proof. The proof follows from the fact that

$$\frac{1}{v(t)}\left(\frac{v^2(t)}{r(t)}\left(\frac{1}{v(t)}\, y'(t)\right)'\right)' = \left(\frac{1}{r(t)}\, y''(t)\right)' + p(t)y'(t). \tag{2.1}$$

□

Now, in the sequel, instead of studying properties of the trinomial equation (E), we will study the behavior of the binomial equation (E^c). For our next considerations, it is desirable for (E^c) to be in a canonical form; that is,

$$\int^{\infty} v(t)\mathrm{d}t = \infty, \tag{2.2}$$

$$\int^{\infty} \frac{r(t)}{v^2(t)}\mathrm{d}t = \infty, \tag{2.3}$$

because properties of the canonical equations are nicely explored.

Now, we will study the properties of the positive solutions of (E_v) to recognize when (2.2)-(2.3) are satisfied. The following result (see, e.g., [7, 9] or [14]) is a consequence of Sturm's comparison theorem.

Lemma 2.2. *If*

$$\frac{R^2(t)}{r(t)}p(t) \le \frac{1}{4}, \tag{2.4}$$

then (E_v) possesses a positive solution $v(t)$.

To be sure that (E_v) possesses a positive solution, we will assume throughout the paper that (2.4) holds. The following result is obvious.

to Contents, page 1

Lemma 2.3. *If $v(t)$ is a positive solution of (E_v), then $v'(t) > 0$, $((1/r(t))v'(t))' < 0$, and, what is more,* (2.2) *holds and there exists $c > 0$ such that $v(t) \le cR(t)$.*

Now, we will show that if (E_v) is nonoscillatory, then we always can choose a positive solution $v(t)$ of (E_v) for which (2.3) holds.

Lemma 2.4. *If $v_1(t)$ is a positive solution of (E_v) for which* (2.3) *is violated, then*

$$v_2(t) = v_1(t) \int_{t_0}^{\infty} \frac{r(s)}{v_1^2(s)} \mathrm{d}s \tag{2.5}$$

is another positive solution of (E_v) and, for $v_2(t)$, (2.3) *holds.*

Proof. First note that

$$v_2''(t) = v_1''(t) \int_{t_0}^{t} \frac{r(s)}{v_1^2(s)} \mathrm{d}s = -p(t)v_1(t) \int_{t_0}^{t} v_1^{-2}(s)\mathrm{d}s = -p(t)v_2(t). \tag{2.6}$$

Thus, $v_2(t)$ is a positive solution of (E_v). On the other hand, to insure that (2.3) holds for $v_2(t)$, let us denote $w(t) = \int_t^\infty r(s)/v_1^2(s)\,\mathrm{d}s$. Then $\lim_{t\to\infty} w(t) = 0$ and

$$\int_{t_1}^{\infty} \frac{r(s)}{v_2^2(s)} \mathrm{d}s = \int_{t_1}^{\infty} \frac{-w'(s)}{w(s)} \mathrm{d}s = \lim_{t\to\infty} \left(\frac{1}{w(t)} - \frac{1}{w(t_1)} \right) = \infty. \tag{2.7}$$

□

Combining Lemmas 2.2, 2.3, and 2.4, we obtain the following result.

Lemma 2.5. *Let* (2.4) *hold. Then trinomial (E) can be represented in its binomial canonical form (E^c).*

Now we can study properties of (E) with help of its canonical representation (E^c). For our reference, let us denote for (E^c)

$$L_0 y = y, \quad L_1 y = \frac{1}{v} (L_0 y)', \quad L_2 y = \frac{v^2}{r} (L_1 y)', \quad L_3 y = (L_2 y)'. \tag{2.8}$$

Now, (E^c) can be written as $L_3 y(t) + v(t)q(t)y(\sigma(t)) = 0$.

We present a structure of the nonoscillatory solutions of (E^c). Since (E^c) is in a canonical form, it follows from the well-known lemma of Kiguradze (see, e.g., [7, 9, 14]) that every nonoscillatory solution $y(t)$ of (E^c) is either of *degree 0*, that is,

$$yL_0 y(t) > 0, \quad yL_1 y(t) < 0, \quad yL_2 y(t) > 0, \quad yL_3 y(t) < 0, \tag{2.9}$$

or of *degree 2*, that is,

$$yL_0 y(t) > 0, \quad yL_1 y(t) > 0, \quad yL_2 y(t) > 0, \quad yL_3 y(t) < 0. \tag{2.10}$$

to Contents, page 1

Definition 2.6. We say that (E^c) has property (A) if its every nonoscillatory solution $y(t)$ is of degree 0; that is, it satisfies (2.9).

Now we verify that property (P_0) of (E) and property (A) of (E^c) are equivalent in the sense that $y(t)$ satisfies (1.1) if and only if it obeys (2.9).

Theorem 2.7. *Let* (2.4) *hold. Assume that $v(t)$ is a positive solution of (E_v) satisfying* (2.2)-(2.3). *Then (E^c) has property (A) if and only if (E) has property (P_0).*

Proof. $\rightarrow$ We suppose that $y(t)$ is a positive solution of (E). We need to verify that $y'(t) < 0$. Since $y(t)$ is also a solution of (E^c), then it satisfies (2.9). Therefore, $0 > L_1y(t) = y'(t)/v(t)$.

$\leftarrow$ Assume that $y(t)$ is a positive solution of (E^c). We will verify that (2.9) holds. Since $y(t)$ is also a solution of (E), we see that $y'(t) < 0$; that is, $L_1y(t) < 0$. It follows from (E^c) that $L_3y(t) = -v(t)q(t)y(\sigma(t)) < 0$. Thus, $L_2y(t)$ is decreasing. If we admit $L_2y(t) < 0$ eventually, then $L_1y(t)$ is decreasing, and integrating the inequality $L_1y(t) < L_1y(t_1)$, we get $y(t) < y(t_1) + L_1y(t_1)\int_{t_1}^{t} v(s)\, ds \rightarrow -\infty$ as $t \rightarrow \infty$. Therefore, $L_2y(t) > 0$ and (2.9) holds. □

The following result which can be found in [9, 14] presents the relationship between property (A) of delay equation and that of equation without delay.

Theorem 2.8. *Let* (2.4) *hold. Assume that $v(t)$ is a positive solution of (E_v) satisfying* (2.2)-(2.3). *Let*

$$\sigma(t) \in C^1([t_0, \infty)), \quad \sigma'(t) > 0. \tag{2.11}$$

If

$$\left(\frac{v^2(t)}{r(t)}\left(\frac{1}{v(t)}\, y'(t)\right)'\right)' + \frac{v(\sigma^{-1}(t))q(\sigma^{-1}(t))}{\sigma'(\sigma^{-1}(t))}\, y(t) = 0 \tag{E_2}$$

has property (A), then so does (E^c).

Combining Theorems 2.7 and 2.8, we get a criterion that reduces property (P_0) of (E) to the property (A) of (E_2).

Corollary 2.9. *Let* (2.4) *and* (2.11) *hold. Assume that $v(t)$ is a positive solution of (E_v) satisfying* (2.2)-(2.3). *If (E_2) has property (A) then (E) has property (P_0).*

Employing any known or future result for property (A) of (E_2), then in view of Corollary 2.9, we immediately obtain that property (P_0) holds for (E).

Example 2.10. We consider the third-order delay trinomial differential equation

$$\left(\frac{1}{t}y''(t)\right)' + \frac{\alpha(2-\alpha)}{t^3}y'(t) + q(t)y(\sigma(t)) = 0, \tag{2.12}$$

to Contents, page 1

where $0 < \alpha < 1$ and $\sigma(t)$ satisfies (2.11). The corresponding equation (E_v) takes the form

$$\left(\frac{1}{t}v'(t)\right)' + \frac{\alpha(2-\alpha)}{t^3}\, v(t) = 0, \tag{2.13}$$

and it has the pair of the solutions $v(t) = t^{\alpha}$ and $\widehat{v}(t) = t^{2-\alpha}$. Thus, $v(t) = t^{\alpha}$ is our desirable solution, which permits to rewrite (2.12) in its canonical form. Then, by Corollary 2.9, (2.12) has property (P_0) if the equation

$$\left(t^{2\alpha-1}\left(t^{-\alpha}\, y'(t)\right)'\right)' + \frac{\left(\sigma^{-1}(t)\right)^{\alpha} q\left(\sigma^{-1}(t)\right)}{\sigma'\left(\sigma^{-1}(t)\right)} y(t) = 0 \tag{2.14}$$

has property (A).

Now, we enhance our results to guarantee stronger asymptotic behavior of the nonoscillatory solutions of (E). We impose an additional condition on the coefficients of (E) to achieve that every nonoscillatory solution of (E) tends to zero as $t \to \infty$.

Corollary 2.11. *Let* (2.4) *and* (2.11) *hold. Assume that $v(t)$ is a positive solution of (E_v) satisfying* (2.2)-(2.3). *If (E_2) has property (A) and*

$$\int_{t_0}^{\infty} v(s_3) \int_{s_3}^{\infty} \frac{r(s_2)}{v^2(s_2)} \int_{s_2}^{\infty} v(s_1) q(s_1) ds_1\, ds_2\, ds_3 = \infty, \tag{2.15}$$

then every nonoscillatory solution $y(t)$ of (E) satisfies (1.2).

Proof. Assume that $y(t)$ is a positive solution of (E). Then, it follows from Corollary 2.9 that $y'(t) < 0$. Therefore, $\lim_{t\to\infty} y(t) = \ell \geq 0$. Assume $\ell > 0$. On the other hand, $y(t)$ is also a solution of (E^c), and, in view of Theorem 2.7, it has to be of *degree 0*; that is, (2.9) is fulfilled. Then, integrating (E^c) from t to ∞, we get

$$L_2 y(t) \geq \int_t^{\infty} v(s) q(s) y(\sigma(s)) \mathrm{d}s \geq \ell \int_t^{\infty} v(s) q(s) \mathrm{d}s. \tag{2.16}$$

Multiplying this inequality by $r(t)/v^2(t)$ and then integrating from t to ∞, we have

$$-L_1 y(t) \geq \ell \int_t^{\infty} \frac{r(s_2)}{v^2(s_2)} \int_{s_2}^{\infty} v(s_1) q(s_1) \mathrm{d}s_1\, \mathrm{d}s_2. \tag{2.17}$$

Multiplying this by $v(t)$ and then integrating from t_1 to t, we obtain

$$y(t_1) \geq \ell \int_{t_1}^{t} v(s_3) \int_{s_3}^{\infty} \frac{r(s_2)}{v^2(s_2)} \int_{s_2}^{\infty} v(s_1) q(s_1) \mathrm{d}s_1\, \mathrm{d}s_2\, \mathrm{d}s_3 \longrightarrow \infty \quad \text{as } t \longrightarrow \infty. \tag{2.18}$$

This is a contradiction, and we deduce that $\ell = 0$. The proof is complete. □

to Contents, page 1

Example 2.12. We consider once more the third-order equation (2.12). It is easy to see that (2.15) takes the form

$$\int_{t_0}^{\infty} s_3^{\alpha} \int_{s_3}^{\infty} s_2^{1-2\alpha} \int_{s_2}^{\infty} s_1^{\alpha} q(s_1) \mathrm{d}s_1 \, \mathrm{d}s_2 \, \mathrm{d}s_3 = \infty. \tag{2.19}$$

Then, by Corollary 2.11, every nonoscillatory solution of (2.12) tends to zero as $t \to \infty$ provided that (2.19) holds and (2.14) has property (A).

In the second part of this paper, we derive criteria that enable us to deduce property (P_0) of (E) from the oscillation of a suitable second-order differential equation. The following theorem is a modification of Tanaka's result [21].

Theorem 2.13. *Let* (2.4) *and* (2.11) *hold. Assume that* $v(t)$ *is a positive solution of* (E_v) *satisfying* (2.2)-(2.3). *Let*

$$\int^{\infty} v(s)q(s)ds < \infty. \tag{2.20}$$

If the second-order equation

$$\left(\frac{v^2(t)}{r(t)} z'(t)\right)' + \left(v(\sigma(t))\sigma'(t) \int_t^{\infty} v(s)q(s)ds\right) z(\sigma(t)) = 0 \tag{E_3}$$

is oscillatory, then (E^c) *has property* (A).

Proof. Assume that $y(t)$ is a positive solution of (E^c), then $y(t)$ is either of *degree 0* or of *degree 2*. Assume that $y(t)$ is of *degree 2*; that is, (2.10) holds. An integration of (E^c) yields

$$L_2 y(t) \geq \int_t^{\infty} v(s)q(s)y(\sigma(s))\mathrm{d}s. \tag{2.21}$$

On the other hand,

$$y(t) \geq \int_{t_1}^{t} v(x)L_1 y(x)\mathrm{d}x. \tag{2.22}$$

Combining the last two inequalities, we get

$$\begin{aligned} L_2 y(t) &\geq \int_t^{\infty} v(s)q(s) \int_{t_1}^{\sigma(s)} v(x)L_1 y(x)\mathrm{d}x \, \mathrm{d}s \\ &\geq \int_t^{\infty} v(s)q(s) \int_{\sigma(t)}^{\sigma(s)} v(x)L_1 y(x)\mathrm{d}x \, \mathrm{d}s \\ &= \int_{\sigma(t)}^{\infty} L_1 y(x)v(x) \int_{\sigma^{-1}(x)}^{\infty} v(s)q(s)\mathrm{d}s \, \mathrm{d}x. \end{aligned} \tag{2.23}$$

to Contents, page 1

Integrating the previous inequality from t_1 to t, we see that $w(t) \equiv L_1y(t)$ satisfies

$$w(t) \geq w(t_1) + \int_{t_1}^{t} \frac{r(s)}{v^2(s)} \int_{\sigma(s)}^{\infty} L_1y(x)v(x) \int_{\sigma^{-1}(x)}^{\infty} v(\delta)q(\delta)\mathrm{d}\delta \,\mathrm{d}x \,\mathrm{d}s. \tag{2.24}$$

Denoting the right-hand side of (2.24) by $z(t)$, it is easy to see that $z(t) > 0$ and

$$\begin{aligned} 0 &= \left(\frac{v^2(t)}{r(t)}z'(t)\right)' + \left(v(\sigma(t))\sigma'(t)\int_t^{\infty} v(s)g(s)ds\right)w(\sigma(t)) = 0 \\ &\geq \left(\frac{v^2(t)}{r(t)}z'(t)\right)' + \left(v(\sigma(t))\sigma'(t)\int_t^{\infty} v(s)g(s)ds\right)z(\sigma(t)) = 0. \end{aligned} \tag{2.25}$$

By Theorem 2 in [14], the corresponding equation (E_3) also has a positive solution. This is a contradiction. We conclude that $y(t)$ is of *degree 0*; that is, (E^c) has property (A). □

If (2.20) does not hold, then we can use the following result.

Theorem 2.14. *Let* (2.4) *and* (2.11) *hold. Assume that* $v(t)$ *is a positive solution of* (E_v) *satisfying* (2.2)-(2.3). *If*

$$\int^{\infty} v(s)q(s)ds = \infty, \tag{2.26}$$

then (E^c) *has property (A).*

Proof. Assume that $y(t)$ is a positive solution of (E^c) and $y(t)$ is of *degree* 2. An integration of (E^c) yields

$$\begin{aligned} L_2y(t_1) &\geq \int_{t_1}^{t} v(s)q(s)y(\sigma(s))\mathrm{d}s \\ &\geq y(\sigma(t_1))\int_{t_1}^{t} v(s)q(s)\mathrm{d}s \longrightarrow \infty \quad \text{as } t \longrightarrow \infty, \end{aligned} \tag{2.27}$$

which is a contradiction. Thus, $y(t)$ is of *degree 0*. The proof is complete now. □

Taking Theorem 2.13 and Corollary 2.9 into account, we get the following criterion for property (P_0) of (E).

Corollary 2.15. *Let* (2.4), (2.11), *and* (2.20) *hold. Assume that* $v(t)$ *is a positive solution of* (E_v) *satisfying* (2.2)-(2.3). *If* (E_3) *is oscillatory, then* (E) *has property* (P_0).

Applying any criterion for oscillation of (E_3), Corollary 2.15 yields a sufficient condition property (P_0) of (E).

to Contents, page 1

Corollary 2.16. *Let* (2.4), (2.11), *and* (2.20) *hold. Assume that* $v(t)$ *is a positive solution of* (E_v) *satisfying* (2.2)-(2.3). *If*

$$\liminf_{t\to\infty}\left(\int_{t_0}^{\sigma(t)}\frac{r(s)}{v^2(s)}ds\right)\left(\int_t^{\infty}v(\sigma(x))\sigma'(x)\int_x^{\infty}v(s)g(s)ds\,dx\right)>\frac{1}{4}, \tag{2.28}$$

then (E) *has property* (P_0).

Proof. It follows from Theorem 11 in [9] that condition (2.28) guarantees the oscillation of (E_3). The proof arises from Corollary 2.16. □

Imposing an additional condition on the coefficients of (E), we can obtain that every nonoscillatory solution of (E) tends to zero as $t\to\infty$.

Corollary 2.17. *Let* (2.4) *and* (2.11) *hold. Assume that* $v(t)$ *is a positive solution of* (E_v) *satisfying* (2.2)-(2.3). *If* (2.28) *and* (2.15) *hold, then every nonoscillatory solution* $y(t)$ *of* (E) *satisfies* (1.2).

Example 2.18. We consider again (2.12). By Corollary 2.17, every nonoscillatory solution of (2.12) tends to zero as $t\to\infty$ provided that (2.19) holds and

$$\liminf_{t\to\infty}\sigma^{2-2\alpha}(t)\left(\int_t^{\infty}\sigma^{\alpha}(x)\sigma'(x)\int_x^{\infty}s^{\alpha}q(s)\mathrm{d}s\,\mathrm{d}x\right)>\frac{2-2\alpha}{4}. \tag{2.29}$$

For a special case of (2.12), namely, for

$$\left(\frac{1}{t}\,y''(t)\right)'+\frac{\alpha(2-\alpha)}{t^3}y'(t)+\frac{a}{t^4}\,y(\lambda t))=0, \tag{2.30}$$

with $0<\alpha<1$, $0<\lambda<1$, and $a>0$, we get that every nonoscillatory solution of (2.30) tends to zero as $t\to\infty$ provided that

$$\frac{a\lambda^{3-\alpha}}{(3-\alpha)(1-\alpha)^2}>1. \tag{2.31}$$

If we set $a=\beta[(\beta+1)(\beta+3)+\alpha(2-\alpha)]\lambda^{\beta}$, where $\beta>0$, then one such solution of (2.12) is $y(t)=t^{-\beta}$.

On the other hand, if for some $\gamma\in(1+\alpha,3-\alpha)$ we have $a=\gamma[(\gamma-1)(3-\gamma)+\alpha(\alpha-2)]\lambda^{-\gamma}>0$, then (2.31) is violated and (2.12) has a nonoscillatory solution $y(t)=t^{\gamma}$ which is of *degree 2*.

3. Summary

In this paper, we have introduced new comparison theorems for the investigation of properties of third-order delay trinomial equations. The comparison principle established in Corollaries 2.9 and 2.11 enables us to deduce properties of the trinomial third-order equations from that of binomial third-order equations. Moreover, the comparison theorems presented in Corollaries 2.15–2.17 permit to derive properties of the trinomial third-order equations from

to Contents, page 1

the oscillation of suitable second-order equations. The results obtained are of high generality, are easily applicable, and are illustrated on suitable examples.

Acknowledgment

This research was supported by S.G.A. KEGA 019-025TUKE-4/2010.

References

[1] B. Baculíková, E. M. Elabbasy, S. H. Saker, and J. Džurina, "Oscillation criteria for third-order nonlinear differential equations," *Mathematica Slovaca*, vol. 58, no. 2, pp. 201–220, 2008.

[2] B. Baculíková, "Oscillation criteria for second order nonlinear differential equations," *Archivum Mathematicum*, vol. 42, no. 2, pp. 141–149, 2006.

[3] B. Baculíková, "Oscillation of certain class of the third order neutral differential equations," *Abstract and Applied Analysis*. In press.

[4] B. Baculíková and D. Lacková, "Oscillation criteria for second order retarded differential equations," *Studies of the University of Žilina. Mathematical Series*, vol. 20, no. 1, pp. 11–18, 2006.

[5] M. Bartušek, M. Cecchi, Z. Došlá, and M. Marini, "Oscillation for third-order nonlinear differential equations with deviating argument," *Abstract and Applied Analysis*, vol. 2010, Article ID 278962, 19 pages, 2010.

[6] R. Bellman, *Stability Theory of Differential Equations*, McGraw-Hill, New York, NY, USA, 1953.

[7] T. A. Chanturija and I. T. Kiguradze, *Asymptotic Properties of Nonautonomous Ordinary Differential Equations*, Nauka, Moscow, Russia, 1990.

[8] J. Džurina, "Asymptotic properties of the third order delay differential equations," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 26, no. 1, pp. 33–39, 1996.

[9] J. Džurina, "Comparison theorems for nonlinear ODEs," *Mathematica Slovaca*, vol. 42, no. 3, pp. 299–315, 1992.

[10] J. Džurina and R. Kotorová, "Properties of the third order trinomial differential equations with delay argument," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 5-6, pp. 1995–2002, 2009.

[11] L. Erbe, "Existence of oscillatory solutions and asymptotic behavior for a class of third order linear differential equations," *Pacific Journal of Mathematics*, vol. 64, no. 2, pp. 369–385, 1976.

[12] P. Hartman, *Ordinary Differential Equations*, John Wiley & Sons, New York, NY, USA, 1964.

[13] G. D. Jones, "An asymptotic property of solutions of $y''' + p(x)y' + q(x)y = 0$," *Pacific Journal of Mathematics*, vol. 47, pp. 135–138, 1973.

[14] T. Kusano and M. Naito, "Comparison theorems for functional-differential equations with deviating arguments," *Journal of the Mathematical Society of Japan*, vol. 33, no. 3, pp. 509–532, 1981.

[15] T. Kusano, M. Naito, and K. Tanaka, "Oscillatory and asymptotic behaviour of solutions of a class of linear ordinary differential equations," *Proceedings of the Royal Society of Edinburgh A*, vol. 90, no. 1-2, pp. 25–40, 1981.

[16] A. C. Lazer, "The behavior of solutions of the differential equation $y''' + p(x)y' + q(x)y = 0$," *Pacific Journal of Mathematics*, vol. 17, pp. 435–466, 1966.

[17] W. E. Mahfoud, "Comparison theorems for delay differential equations," *Pacific Journal of Mathematics*, vol. 83, no. 1, pp. 187–197, 1979.

[18] N. Parhi and S. Padhi, "On asymptotic behavior of delay-differential equations of third order," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 34, no. 3, pp. 391–403, 1998.

[19] N. Parhi and S. Padhi, "Asymptotic behaviour of solutions of third order delay-differential equations," *Indian Journal of Pure and Applied Mathematics*, vol. 33, no. 10, pp. 1609–1620, 2002.

[20] A. Škerlík, "Integral criteria of oscillation for a third order linear differential equation," *Mathematica Slovaca*, vol. 45, no. 4, pp. 403–412, 1995.

[21] K. Tanaka, "Asymptotic analysis of odd order ordinary differential equations," *Hiroshima Mathematical Journal*, vol. 10, no. 2, pp. 391–408, 1980.

[22] W. F. Trench, "Canonical forms and principal systems for general disconjugate equations," *Transactions of the American Mathematical Society*, vol. 189, pp. 319–327, 1973.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 709427, 15 pages
doi:10.1155/2011/709427

Research Article

Asymptotic Convergence of the Solutions of a Discrete Equation with Two Delays in the Critical Case

L. Berezansky,[1] J. Diblík,[2] M. Růžičková,[2] and Z. Šutá[2]

[1] *Department of Mathematics, Ben-Gurion University of the Negev, 84105 Beer-Sheva, Israel*
[2] *Department of Mathematics, University of Žilina, 01026 Žilina, Slovakia*

Correspondence should be addressed to J. Diblík, diblik@feec.vutbr.cz

Received 9 October 2010; Revised 17 March 2011; Accepted 13 April 2011

Academic Editor: Elena Braverman

A discrete equation $\Delta y(n) = \beta(n)[y(n-j) - y(n-k)]$ with two integer delays k and j, $k > j \geq 0$ is considered for $n \to \infty$. We assume $\beta : \mathbb{Z}_{n_0-k}^{\infty} \to (0,\infty)$, where $\mathbb{Z}_{n_0}^{\infty} = \{n_0, n_0+1, \ldots\}$, $n_0 \in \mathbb{N}$ and $n \in \mathbb{Z}_{n_0}^{\infty}$. Criteria for the existence of strictly monotone and asymptotically convergent solutions for $n \to \infty$ are presented in terms of inequalities for the function β. Results are sharp in the sense that the criteria are valid even for some functions β with a behavior near the so-called critical value, defined by the constant $(k-j)^{-1}$. Among others, it is proved that, for the asymptotic convergence of all solutions, the existence of a strictly monotone and asymptotically convergent solution is sufficient.

1. Introduction

We use the following notation: for integers $s, q, s \leq q$, we define $\mathbb{Z}_s^q := \{s, s+1, \ldots, q\}$, where the cases $s = -\infty$ and $q = \infty$ are admitted too. Throughout this paper, using the notation $\mathbb{Z}_s^q$ or another one with a pair of integers s, q, we assume $s \leq q$.

In this paper we study a discrete equation with two delays

$$\Delta y(n) = \beta(n)\left[y(n-j) - y(n-k)\right] \tag{1.1}$$

as $n \to \infty$. Integers k and j in (1.1) satisfy the inequality $k > j \geq 0$ and $\beta : \mathbb{Z}_{n_0-k}^{\infty} \to \mathbb{R}^+ := (0,\infty)$, where $n_0 \in \mathbb{N}$ and $n \in \mathbb{Z}_{n_0}^{\infty}$. Without loss of generality, we assume $n_0 - k > 0$ throughout the paper (this is a technical detail, necessary for some expressions to be well defined).

to Contents, page 1

The results concern the asymptotic convergence of all solutions of (1.1). We focus on what is called the *critical case* (with respect to the function β) which separates the case when all solutions are convergent from the case when there exist divergent solutions.

Such a critical case is characterized by the constant value

$$\beta(n) \equiv \beta_{\text{cr}} := (k-j)^{-1}, \quad n \in \mathbb{Z}_{n_0-k}^{\infty}, \tag{1.2}$$

and below we explain its meaning and importance by an analysis of the asymptotic behavior of solutions of (1.1).

Consider (1.1) with $\beta(n) = \beta_0$, where β_0 is a positive constant; that is, we consider the following equation:

$$\Delta y(n) = \beta_0 \cdot [y(n-j) - y(n-k)]. \tag{1.3}$$

Looking for a solution of (1.3) in the form $y(n) = \lambda^n$, $\lambda \in \mathbb{C} \setminus \{0\}$ using the usual procedure, we get the characteristic equation

$$\lambda^{k+1} - \lambda^k = \beta_0 \cdot \left[\lambda^{k-j} - 1\right]. \tag{1.4}$$

Denote its roots by λ_i, $i = 1, \dots, k+1$. Then characteristic equation (1.4) has a root $\lambda_{k+1} = 1$. Related solution of (1.3) is $y_{k+1}(n) = 1$. Then there exists a one-parametric family of constant solutions of (1.3) $y(n) = c_{k+1}y_{k+1}(n) = c_{k+1}$, where c_{k+1} is an arbitrary constant. Equation (1.4) can be rewritten as

$$\lambda^k(\lambda - 1) = \beta_0 \cdot (\lambda - 1)\left(\lambda^{k-j-1} + \lambda^{k-j-2} + \cdots + 1\right), \tag{1.5}$$

and, instead of (1.4), we can consider the following equation:

$$f(\lambda) := \lambda^k - \beta_0 \cdot \left(\lambda^{k-j-1} + \lambda^{k-j-2} + \cdots + 1\right) = 0. \tag{1.6}$$

Let $\beta_0 = \beta_{\text{cr}}$. Then (1.6) has a root $\lambda_k = 1$ which is a double root of (1.4). By the theory of linear difference equations, (1.3) has a solution $y_k(n) = n$, linearly independent with $y_{k+1}(n)$. There exists a two-parametric family of solutions of (1.3)

$$y(n) = c_k y_k(n) + c_{k+1}y_{k+1}(n) = c_k n + c_{k+1}, \tag{1.7}$$

where c_k, c_{k+1} are arbitrary constants. Then $\lim_{n\to\infty} y(n) = \infty$ if $c_k \neq 0$. This means that solutions with $c_k \neq 0$ are divergent.

Let $\beta_0 < \beta_{\text{cr}}$ and $k - j > 1$. We define two functions of a complex variable λ

$$F(\lambda) := \lambda^k, \qquad \Psi(\lambda) := \beta_0 \cdot \left(\lambda^{k-j-1} + \lambda^{k-j-2} + \cdots + 1\right), \tag{1.8}$$

to Contents, page 1

and (1.6) can be written as

$$F(\lambda) - \Psi(\lambda) = 0. \tag{1.9}$$

By Rouche's theorem, all roots λ_i, $i = 1, 2, \ldots, k$ of (1.6) satisfy $|\lambda_i| < 1$ because, on the boundary C of a unit circle $|\lambda| < 1$, we have

$$|\Psi(\lambda)|_C = \beta_0 \cdot \left|\lambda^{k-j-1} + \lambda^{k-j-2} + \cdots + 1\right| < \frac{1}{k-j}(k-j) = 1 = |F(\lambda)|_C, \tag{1.10}$$

and the functions $F(\lambda)$, $F(\lambda) - \Psi(\lambda)$ have the same number of zeros in the domain $|\lambda| < 1$.

The case $\beta_0 < \beta_{\text{cr}}$ and $k - j = 1$ is trivial because (1.6) turns into

$$\lambda^k - \beta_0 = 0 \tag{1.11}$$

and, due to inequality $|\lambda|^k = \beta_0 < \beta_{\text{cr}} = 1$, has all its roots in the domain $|\lambda| < 1$.

Then the relevant solutions $y_i(n)$, $i = 1, 2, \ldots, k$ satisfy $\lim_{n\to\infty} y_i(n) = 0$, and the limit of the general solution of (1.3), $y(n) = \lim_{n\to\infty} \sum_{i=1}^{k+1} c_i y_i(n)$ where c_i are arbitrary constants, is finite because

$$\lim_{n\to\infty} y(n) = \lim_{n\to\infty} \sum_{i=1}^{k+1} c_i y_i(n) = c_{k+1}. \tag{1.12}$$

Let $\beta_0 > \beta_{\text{cr}}$. Since $f(1) = 1 - \beta_0 \cdot (k-j) < 0$ and $f(+\infty) = +\infty$, there exists a root $\lambda = \lambda_* > 1$ of (1.6) and a solution $y_*(n) = (\lambda_*)^n$ of (1.3) satisfying $\lim_{n\to\infty} y_*(n) = \infty$. This means that solution $y_*(n)$ is divergent.

Gathering all the cases considered, we have the following:

(i) if $0 < \beta_0 < \beta_{\text{cr}}$, then all solutions of (1.3) have a finite limit as $n \to \infty$,

(ii) if $\beta_0 \geq \beta_{\text{cr}}$, then there exists a divergent solution of (1.3) when $n \to \infty$.

The above analysis is not applicable in the case of a nonconstant function $\beta(n)$ in (1.1). To overcome some difficulties, the method of auxiliary inequalities is applied to investigate (1.1). From our results it follows that, for example, all solutions of (1.1) have a finite limit for $n \to \infty$ (or, in accordance with the below definition, are asymptotically convergent) if there exists a $p > 1$ such that the inequality

$$\beta(n) \leq \frac{1}{k-j} - \frac{p(k+j+1)}{2n(k-j)} \tag{1.13}$$

holds for all $n \in \mathbb{Z}^{\infty}_{n_0-k}$, where n_0 is a sufficiently large natural number. The limit of the right-hand side of (1.13) as $n \to \infty$ equals the critical value β_{cr}:

$$\lim_{n\to\infty}\left(\frac{1}{k-j} - \frac{p(k+j+1)}{2n(k-j)}\right) = \frac{1}{k-j} = \beta_{\text{cr}}. \tag{1.14}$$

to Contents, page 1

It means that the function $\beta(n)$ in (1.1) can be sufficiently close to the critical value β_{cr} but such that all solutions of (1.1) are convergent as $n \to \infty$.

The proofs of the results are based on comparing the solutions of (1.1) with those of an auxiliary inequality that formally copies (1.1). First, we prove that, under certain conditions, (1.1) has an increasing and convergent solution $y = y(n)$ (i.e., there exists a finite limit $\lim_{n\to\infty} y(n)$). Then we extend this statement to all the solutions of (1.1). It is an interesting fact that, in the general case, the asymptotic convergence of all solutions is characterized by the existence of a strictly increasing and bounded solution.

The problem concerning the asymptotic convergence of solutions in the continuous case, that is, in the case of delayed differential equations or other classes of equations, is a classical one and has attracted much attention recently. The problem of the asymptotic convergence of solutions of discrete and difference equations with delay has not yet received much attention. We mention some papers from both of these fields (in most of them, equations and systems with a structure similar to the discrete equation (1.1) are considered).

Arino and Pituk [1], for example, investigate linear and nonlinear perturbations of a linear autonomous functional-differential equation which has infinitely many equilibria. Bereketoğlu and Karakoç [2] derive sufficient conditions for the asymptotic constancy and estimates of the limits of solutions for an impulsive system, and Györi et al. give sufficient conditions for the convergence of solutions of a nonhomogeneous linear system of impulsive delay differential equations and a limit formula in [3]. Bereketoğlu and Pituk [4] give sufficient conditions for the asymptotic constancy of solutions of nonhomogeneous linear delay differential equations with unbounded delay. The limits of the solutions can be computed in terms of the initial conditions and a special matrix solution of the corresponding adjoint equation. In [5] Diblík studies the scalar equation under the assumption that every constant is its solution. Criteria and sufficient conditions for the convergence of solutions are found. The paper by Diblík and Růžičková [6] deals with the asymptotic behavior of a first-order linear homogeneous differential equation with double delay. The convergence of solutions of the delay Volterra equation in the critical case is studied by Messina et al. in [7]. Berezansky and Braverman study a behavior of solutions of a food-limited population model with time delay in [8].

Bereketoğlu and Huseynov [9] give sufficient conditions for the asymptotic constancy of the solutions of a system of linear difference equations with delays. The limits of the solutions, as $t \to \infty$, can be computed in terms of the initial function and a special matrix solution of the corresponding adjoint equation. Dehghan and Douraki [10] study the global behavior of a certain difference equation and show, for example, that zero is always an equilibrium point which satisfies a necessary and suffient condition for its local asymptotic stability. Györi and Horváth [11] study a system of linear delay difference equations such that every solution has a finite limit at infinity. The stability of difference equations is studied intensively in papers by Stević [12, 13]. In [12], for example, he proves the global asymptotic stability of a class of difference equations. Baštinec and Diblík [14] study a class of positive and vanishing at infinity solutions of a linear difference equation with delay. Nonoscillatory solutions of second-order difference equations of the Poincaré type are investigated by Medina and Pituk in [15].

Comparing the known investigations with the results presented, we can see that our results can be applied to the critical case giving strong sufficient conditions of asymptotic convergence of solutions for this case. Nevertheless, we are not concerned with computing the limits of the solutions as $n \to \infty$.

to Contents, page 1

The paper is organized as follows. In Section 2 auxiliary results are collected, an auxiliary inequality is studied, and the relationship of its solutions with the solutions of (1.1) is derived. The existence of a strictly increasing and convergent solution of (1.1) is established in Section 3. Section 4 contains results concerning the convergence of all solutions of (1.1). An example illustrating the sharpness of the results derived is given as well.

Throughout the paper we adopt the customary notation $\sum_{i=k+s}^{k} \mathcal{B}(i) = 0$, where k is an integer, s is a positive integer, and $\mathcal{B}$ denotes the function under consideration regardless of whether it is defined for the arguments indicated or not.

2. Auxiliary Results

Let $\mathcal{C} := \mathcal{C}(\mathbb{Z}_{-k}^{0}, \mathbb{R})$ be the space of discrete functions mapping the discrete interval $\mathbb{Z}_{-k}^{0}$ into $\mathbb{R}$. Let $v \in \mathbb{Z}_{n_0}^{\infty}$ be given. The function $y : \mathbb{Z}_{v-k}^{\infty} \to \mathbb{R}$ is said to be a *solution of* (1.1) *on* $\mathbb{Z}_{v-k}^{\infty}$ if it satisfies (1.1) for every $n \in \mathbb{Z}_{v}^{\infty}$. A solution y of (1.1) on $\mathbb{Z}_{v-k}^{\infty}$ is *asymptotically convergent* if the limit $\lim_{n\to\infty} y(n)$ exists and is finite. For a given $v \in \mathbb{Z}_{n_0}^{\infty}$ and $\varphi \in \mathcal{C}$, we say that $y = y_{(v,\varphi)}$ is a *solution of* (1.1) defined by the initial conditions (v, φ) if $y_{(v,\varphi)}$ is a solution of (1.1) on $\mathbb{Z}_{v-k}^{\infty}$ and $y_{(v,\varphi)}(v+m) = \varphi(m)$ for $m \in \mathbb{Z}_{-k}^{0}$.

2.1. Auxiliary Inequality

The auxiliary inequality

$$\Delta\omega(n) \geq \beta(n)\left[\omega(n-j) - \omega(n-k)\right] \tag{2.1}$$

will serve as a helpful tool in the analysis of (1.1). Let $v \in \mathbb{Z}_{n_0}^{\infty}$. The function $\omega : \mathbb{Z}_{v-k}^{\infty} \to \mathbb{R}$ is said to be a *solution* of (2.1) on $\mathbb{Z}_{v-k}^{\infty}$ if ω satisfies inequality (2.1) for $n \in \mathbb{Z}_{v}^{\infty}$. A solution ω of (2.1) on $\mathbb{Z}_{v-k}^{\infty}$ is *asymptotically convergent* if the limit $\lim_{n\to\infty}\omega(n)$ exists and is finite.

We give some properties of solutions of inequalities of the type (2.1), which will be utilized later on. We will also compare the solutions of (1.1) with the solutions of inequality (2.1).

Lemma 2.1. *Let $\varphi \in \mathcal{C}$ be strictly increasing (nondecreasing, strictly decreasing, nonincreasing) on $\mathbb{Z}_{-k}^{0}$. Then the corresponding solution $y_{(n^*,\varphi)}(n)$ of* (1.1) *with $n^* \in \mathbb{Z}_{n_0}^{\infty}$ is strictly increasing (nondecreasing, strictly decreasing, nonincreasing) on $\mathbb{Z}_{n^*-k}^{\infty}$ too.*

If φ is strictly increasing (nondecreasing) and $\omega : \mathbb{Z}_{n_0-k}^{\infty} \to \mathbb{R}$ is a solution of inequality (2.1) *with $\omega(n_0+m) = \varphi(m)$, $m \in \mathbb{Z}_{n_0-k}^{n_0}$, then ω is strictly increasing (nondecreasing) on $\mathbb{Z}_{n_0-k}^{\infty}$.*

Proof. This follows directly from (1.1), inequality (2.1), and from the properties $\beta(n) > 0$, $n \in \mathbb{Z}_{n_0-k}^{\infty}$, $k > j \geq 0$. □

Theorem 2.2. *Let $\omega(n)$ be a solution of inequality* (2.1) *on $\mathbb{Z}_{n_0-k}^{\infty}$. Then there exists a solution $y(n)$ of* (1.1) *on $\mathbb{Z}_{n_0-k}^{\infty}$ such that the inequality*

$$y(n) \leq \omega(n) \tag{2.2}$$

to Contents, page 1

holds on $\mathbb{Z}_{n_0-k}^{\infty}$. *In particular, a solution* $y(n_0,\phi)$ *of* (1.1) *with* $\phi \in \mathcal{C}$ *defined by the equation*

$$\phi(m) := \omega(n_0+m), \quad m \in \mathbb{Z}_{-k}^{0} \tag{2.3}$$

is such a solution.

Proof. Let $\omega(n)$ be a solution of inequality (2.1) on $\mathbb{Z}_{n_0-k}^{\infty}$. We will show that the solution $y(n) := y_{(n_0,\phi)}(n)$ of (1.1) satisfies inequality (2.2), that is,

$$y_{(n_0,\phi)}(n) \le \omega(n) \tag{2.4}$$

on $\mathbb{Z}_{n_0-k}^{\infty}$. Let $W : \mathbb{Z}_{n_0-k}^{\infty} \to \mathbb{R}$ be defined by $W(n) = \omega(n) - y(n)$. Then $W = 0$ on $\mathbb{Z}_{n_0-k}^{n_0}$, and, in addition, W is a solution of (2.1) on $\mathbb{Z}_{n_0-k}^{\infty}$. Lemma 2.1 implies that W is nondecreasing. Consequently, $\omega(n) - y(n) \ge \omega(n_0) - y(n_0) = 0$ for all $n \ge n_0$. □

2.2. Comparison Lemma

Now we consider an inequality of the type (2.1)

$$\Delta\omega^*(n) \ge \beta_1(n)\left[\omega^*(n-j) - \omega^*(n-k)\right], \tag{2.5}$$

where $\beta_1 : \mathbb{Z}_{n_0-k}^{\infty} \to \mathbb{R}^+$ is a discrete function satisfying $\beta_1(n) \ge \beta(n)$ on $\mathbb{Z}_{n_0-k}^{\infty}$. The following comparison lemma holds.

Lemma 2.3. *Let* $\omega^* : \mathbb{Z}_{n_0-k}^{\infty} \to \mathbb{R}^+$ *be a nondecreasing positive solution of inequality* (2.5) *on* $\mathbb{Z}_{n_0-k}^{\infty}$. *Then* ω^* *is a solution of inequality* (2.1) *on* $\mathbb{Z}_{n_0-k}^{\infty}$ *too.*

Proof. Let ω^* be a nondecreasing solution of (2.5) on $\mathbb{Z}_{n_0-k}^{\infty}$. We have

$$\omega^*(n-j) - \omega^*(n-k) \ge 0 \tag{2.6}$$

because $n-k < n-j$. Then

$$\Delta\omega^*(n) \ge \beta_1(n)\left[\omega^*(n-j) - \omega^*(n-k)\right] \ge \beta(n)\left[\omega^*(n-j) - \omega^*(n-k)\right] \tag{2.7}$$

on $\mathbb{Z}_{n_0}^{\infty}$. Consequently, the function $\omega := \omega^*$ solves inequality (2.1) on $\mathbb{Z}_{n_0}^{\infty}$, too. □

2.3. A Solution of Inequality (2.1)

We will construct a solution of inequality (2.1). In the following lemma, we obtain a solution of inequality (2.1) in the form of a sum. This auxiliary result will help us derive sufficient conditions for the existence of a strictly increasing and asymptotically convergent solution of (1.1) (see Theorem 3.2 below).

to Contents, page 1

Lemma 2.4. *Let there exist a discrete function* $\varepsilon : \mathbb{Z}_{n_0-k}^{\infty} \to \mathbb{R}^+$ *such that*

$$\varepsilon(n+1) \geq \sum_{i=n-k+1}^{n-j} \beta(i-1)\varepsilon(i) \tag{2.8}$$

on $\mathbb{Z}_{n_0}^{\infty}$*. Then there exists a solution* $\omega(n) = \omega_\varepsilon(n)$ *of inequality* (2.1) *defined on* $\mathbb{Z}_{n_0-k}^{\infty}$ *having the form*

$$\omega_\varepsilon(n) := \sum_{i=n_0-k+1}^{n} \beta(i-1)\varepsilon(i). \tag{2.9}$$

Proof. For $n \in \mathbb{Z}_{n_0}^{\infty}$, we get

$$\begin{aligned}
\Delta\omega_\varepsilon(n) &= \omega_\varepsilon(n+1) - \omega_\varepsilon(n) \\
&= \sum_{i=n_0-k+1}^{n+1} \beta(i-1)\varepsilon(i) - \sum_{i=n_0-k+1}^{n} \beta(i-1)\varepsilon(i) \\
&= \beta(n)\varepsilon(n+1), \\
\omega_\varepsilon(n-j) - \omega_\varepsilon(n-k) &= \sum_{i=n_0-k+1}^{n-j} \beta(i-1)\varepsilon(i) - \sum_{i=n_0-k+1}^{n-k} \beta(i-1)\varepsilon(i) \\
&= \sum_{i=n-k+1}^{n-j} \beta(i-1)\varepsilon(i).
\end{aligned} \tag{2.10}$$

We substitute ω_ε for ω in (2.1). Using (2.10), we get

$$\beta(n)\varepsilon(n+1) \geq \beta(n) \sum_{n-k+1}^{n-j} \beta(i-1)\varepsilon(i). \tag{2.11}$$

This inequality will be satisfied if inequality (2.8) holds. Indeed, reducing the last inequality by $\beta(n)$, we obtain

$$\varepsilon(n+1) \geq \sum_{n-k+1}^{n-j} \beta(i-1)\varepsilon(i), \tag{2.12}$$

which is inequality (2.8). □

to Contents, page 1

2.4. Decomposition of a Function into the Difference of Two Strictly Increasing Functions

It is well known that every absolutely continuous function is representable as the difference of two increasing absolutely continuous functions [16, page 318]. We will need a simple discrete analogue of this result.

Lemma 2.5. *Every function $\varphi \in \mathcal{C}$ can be decomposed into the difference of two strictly increasing functions $\varphi_j \in \mathcal{C}$, $j = 1, 2$, that is,*

$$\varphi(n) = \varphi_1(n) - \varphi_2(n), \qquad n \in \mathbb{Z}_{-k}^{0}. \tag{2.13}$$

Proof. Let constants $M_n > 0$, $n \in \mathbb{Z}_{-k}^{0}$ be such that inequalities

$$M_{n+1} > M_n + \max\{0, \varphi(n) - \varphi(n+1)\} \tag{2.14}$$

are valid for $n \in \mathbb{Z}_{-k}^{-1}$. We set

$$\begin{aligned} \varphi_1(n) &:= \varphi(n) + M_n, \quad n \in \mathbb{Z}_{-k}^{0}, \\ \varphi_2(n) &:= M_n, \quad n \in \mathbb{Z}_{-k}^{0}. \end{aligned} \tag{2.15}$$

It is obvious that (2.13) holds. Now we verify that both functions φ_j, $j = 1, 2$ are strictly increasing. The first one should satisfy $\varphi_1(n+1) > \varphi_1(n)$ for $n \in \mathbb{Z}_{-k}^{-1}$, which means that

$$\varphi(n+1) + M_{n+1} > \varphi(n) + M_n \tag{2.16}$$

or

$$M_{n+1} > M_n + \varphi(n) - \varphi(n+1). \tag{2.17}$$

We conclude that the last inequality holds because, due to (2.14), we have

$$M_{n+1} > M_n + \max\{0, \varphi(n) - \varphi(n+1)\} \geq M_n + \varphi(n) - \varphi(n+1). \tag{2.18}$$

The inequality $\varphi_2(n+1) > \varphi_2(n)$ obviously holds for $n \in \mathbb{Z}_{-k}^{-1}$ due to (2.14) as well. □

2.5. Auxiliary Asymptotic Decomposition

The following lemma can be proved easily by induction. The symbol $\mathcal{O}$ stands for the Landau order symbol.

to Contents, page 1

Lemma 2.6. *For fixed* $r, \sigma \in \mathbb{R} \setminus \{0\}$, *the asymptotic representation*

$$(n-r)^{\sigma} = n^{\sigma}\left[1 - \frac{\sigma r}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)\right] \tag{2.19}$$

holds for $n \to \infty$.

3. Convergent Solutions of (1.1)

This part deals with the problem of detecting the existence of asymptotically convergent solutions. The results shown below provide sufficient conditions for the existence of an asymptotically convergent solution of (1.1). First we present two obvious statements concerning asymptotic convergence. From Lemma 2.1 and Theorem 2.2, we immediately get the following.

Theorem 3.1. *Let* $\omega(n)$ *be a strictly increasing and bounded solution of* (2.1) *on* $\mathbb{Z}_{n_0-k}^{\infty}$. *Then there exists a strictly increasing and asymptotically convergent solution* $y(n)$ *of* (1.1) *on* $\mathbb{Z}_{n_0-k}^{\infty}$.

From Lemma 2.1, Theorem 2.2, and Lemma 2.4, we get the following.

Theorem 3.2. *Let there exist a function* $\varepsilon : \mathbb{Z}_{n_0-k}^{\infty} \to \mathbb{R}^{+}$ *satisfying*

$$\sum_{i=n_0-k+1}^{\infty} \beta(i-1)\varepsilon(i) < \infty \tag{3.1}$$

and inequality (2.8) *on* $\mathbb{Z}_{n_0}^{\infty}$. *Then the initial function*

$$\varphi(n) = \sum_{i=n_0-k+1}^{n_0+n} \beta(i-1)\varepsilon(i), \quad n \in \mathbb{Z}_{-k}^{0} \tag{3.2}$$

defines a strictly increasing and asymptotically convergent solution $y_{(n_0,\varphi)}(n)$ *of* (1.1) *on* $\mathbb{Z}_{n_0-k}^{\infty}$ *satisfying the inequality*

$$y(n) \le \sum_{i=n_0-k+1}^{n} \beta(i-1)\varepsilon(i) \tag{3.3}$$

on $\mathbb{Z}_{n_0}^{\infty}$.

Assuming that the coefficient $\beta(n)$ in (1.1) can be estimated by a suitable function, we can prove that (1.1) has a convergent solution.

Theorem 3.3. *Let there exist a* $p > 1$ *such that the inequality*

$$\beta(n) \le \frac{1}{k-j} - \frac{p(k+j+1)}{2n(k-j)} \tag{3.4}$$

to Contents, page 1

holds for all $n \in \mathbb{Z}_{n_0-k}^{\infty}$. *Then there exists a strictly increasing and asymptotically convergent solution* $y(n)$ *of* (1.1) *as* $n \to \infty$.

Proof. In the proof, we assume (without loss of generality) that n_0 is sufficiently large for asymptotic computations to be valid. Let us verify that inequality (2.8) has a solution ε such that

$$\sum_{i=n_0-k+1}^{\infty} \beta(i-1)\varepsilon(i) < \infty. \tag{3.5}$$

We put

$$\beta(n) = \beta^*(n) := \frac{1}{k-j} - \frac{p^*}{2n}, \qquad \varepsilon(n) := \frac{1}{n^\alpha} \tag{3.6}$$

in inequality (2.8), where $p^* > 0$ and $\alpha > 1$ are constants. Then, for the right-hand side $\mathcal{R}(n)$ of (2.8), we have

$$\begin{aligned}
\mathcal{R}(n) &= \sum_{i=n-k+1}^{n-j} \left[\frac{1}{k-j} - \frac{p^*}{2(i-1)}\right]\frac{1}{i^\alpha} \\
&= \frac{1}{k-j}\sum_{i=n-k+1}^{n-j}\frac{1}{i^\alpha} - \frac{p^*}{2}\sum_{i=n-k+1}^{n-j}\frac{1}{(i-1)i^\alpha} \\
&= \frac{1}{k-j}\left[\frac{1}{(n-k+1)^\alpha} + \frac{1}{(n-k+2)^\alpha} + \cdots + \frac{1}{(n-j)^\alpha}\right] \\
&\quad - \frac{p^*}{2}\left[\frac{1}{(n-k)(n-k+1)^\alpha} + \frac{1}{(n-k+1)(n-k+2)^\alpha} + \cdots + \frac{1}{(n-j-1)(n-j)^\alpha}\right].
\end{aligned} \tag{3.7}$$

We asymptotically decompose $\mathcal{R}(n)$ as $n \to \infty$ using decomposition formula (2.19) in Lemma 2.6. We apply this formula to each term in the first square bracket with $\sigma = -\alpha$ and with $r = k-1$ for the first term, $r = k-2$ for the second term, and so forth, and, finally, $r = j$ for the last term. To estimate the terms in the second square bracket, we need only the first terms of the decomposition and the order of accuracy, which can be computed easily without using Lemma 2.6. We get

$$\begin{aligned}
\mathcal{R}(n) = {}& \frac{1}{(k-j)n^\alpha}\left[1 + \frac{\alpha(k-1)}{n} + 1 + \frac{\alpha(k-2)}{n} + \cdots + 1 + \frac{\alpha j}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)\right] \\
& - \frac{p^*}{2n^{\alpha+1}}\left[1 + 1 + \cdots + 1 + \mathcal{O}\left(\frac{1}{n}\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{(k-j)n^{\alpha+1}}\left[(k-j)n + \alpha(k-1) + \alpha(k-2) + \cdots + \alpha j + \mathcal{O}\left(\frac{1}{n}\right)\right] \\
&\quad - \frac{p^*}{2n^{\alpha+1}}\left[(k-j) + \mathcal{O}\left(\frac{1}{n}\right)\right] \\
&= \frac{1}{n^\alpha} + \frac{\alpha}{(k-j)n^{\alpha+1}}\frac{(k+j-1)(k-j)}{2} - \frac{p^*}{2n^{\alpha+1}}(k-j) + \mathcal{O}\left(\frac{1}{n^{\alpha+2}}\right),
\end{aligned}
\tag{3.8}
$$

and, finally,

$$
\mathcal{R}(n) = \frac{1}{n^\alpha} + \frac{\alpha}{2n^{\alpha+1}}(k+j-1) - \frac{p^*}{2n^{\alpha+1}}(k-j) + \mathcal{O}\left(\frac{1}{n^{\alpha+2}}\right). \tag{3.9}
$$

A similar decomposition of the left-hand side $\mathcal{L}(n) = \varepsilon(n+1) = (n+1)^{-\alpha}$ in inequality (2.8) leads to

$$
\mathcal{L}(n) \equiv \frac{1}{(n+1)^\alpha} = \frac{1}{n^\alpha}\left[1 - \frac{\alpha}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)\right] = \frac{1}{n^\alpha} - \frac{\alpha}{n^{\alpha+1}} + \mathcal{O}\left(\frac{1}{n^{\alpha+2}}\right) \tag{3.10}
$$

(we use decomposition formula (2.19) in Lemma 2.6 with $\sigma = -\alpha$ and $r = -1$).

Comparing $\mathcal{L}(n)$ and $\mathcal{R}(n)$, we see that, for $\mathcal{L}(n) \geq \mathcal{R}(n)$, it is necessary to match the coefficients of the terms $n^{-\alpha-1}$ because the coefficients of the terms $n^{-\alpha}$ are equal. It means that we need the inequality

$$
-\alpha > \frac{\alpha(k+j-1)}{2} - \frac{p^*}{2}(k-j). \tag{3.11}
$$

Simplifying this inequality, we get

$$
\begin{aligned}
&\frac{p^*}{2}(k-j) > \alpha + \frac{\alpha(k+j-1)}{2}, \\
&p^*(k-j) > \alpha(k+j+1),
\end{aligned}
\tag{3.12}
$$

and, finally,

$$
p^* > \frac{\alpha(k+j+1)}{k-j}. \tag{3.13}
$$

We set

$$
p^* := p\,\frac{k+j+1}{k-j}, \tag{3.14}
$$

to Contents, page 1

where $p = \text{const}$. Then the previous inequality holds for $p > \alpha$, that is, for $p > 1$. Consequently, the function β^* defined by (3.6) has the form

$$\beta^*(n) = \frac{1}{k-j} - \frac{p(k+j+1)}{2(k-j)n} \tag{3.15}$$

with $p > 1$, and, for the function ω_ε defined by formula (2.9), we have

$$\omega_\varepsilon(n) = \sum_{i=n_0-k+1}^{n} \left(\frac{1}{k-j} - \frac{p(k+j+1)}{2(k-j)(i-1)} \right) \frac{1}{i^\alpha}. \tag{3.16}$$

Function $\omega_\varepsilon(n)$ is a positive solution of inequality (2.1), and, moreover, it is easy to verify that $\omega_\varepsilon(\infty) < \infty$ since $\alpha > 1$. This is a solution to every inequality of the type (2.1) if the function β^* fixed by formula (3.15) is changed by an arbitrary function β satisfying inequality (3.4). This is a straightforward consequence of Lemma 2.3 if, in its formulation, we set

$$\beta_1(n) := \beta^*(n) = \frac{1}{k-j} - \frac{p(k+j+1)}{2(k-j)n} \tag{3.17}$$

with $p > 1$ since $\omega^* \equiv \omega_\varepsilon$ is the desired solution of inequality (2.5). Finally, by Theorem 3.1 with $\omega := \omega_\varepsilon$ as defined by (3.16), we conclude that there exists a strictly increasing and convergent solution $y(n)$ of (1.1) as $n \to \infty$ satisfying the inequality

$$y(n) < \omega_\varepsilon(n) \tag{3.18}$$

on $\mathbb{Z}_{n_0-k}^{\infty}$. □

4. Convergence of All Solutions

In this part we present results concerning the convergence of all solutions of (1.1). First we use inequality (3.4) to state the convergence of all the solutions.

Theorem 4.1. *Let there exist a $p > 1$ such that inequality* (3.4) *holds for all $n \in \mathbb{Z}_{n_0-k}^{\infty}$. Then all solutions of* (1.1) *are convergent as $n \to \infty$.*

Proof. First we prove that every solution defined by a monotone initial function is convergent. We will assume that a strictly monotone initial function $\varphi \in \mathcal{C}$ is given. For definiteness, let φ be strictly increasing or nondecreasing (the case when it is strictly decreasing or nonincreasing can be considered in much the same way). By Lemma 2.1, the solution $y_{(n_0,\varphi)}$ is monotone; that is, it is either strictly increasing or nondecreasing. We prove that $y_{(n_0,\varphi)}$ is convergent.

By Theorem 3.3 there exists a strictly increasing and asymptotically convergent solution $y = Y(n)$ of (1.1) on $\mathbb{Z}_{n_0-k}^{\infty}$. Without loss of generality we assume $y_{(n_0,\varphi)} \not\equiv Y(n)$ on

to Contents, page 1

$\mathbb{Z}_{n_0-k}^{\infty}$ since, in the opposite case, we can choose another initial function. Similarly, without loss of generality, we can assume

$$\Delta Y(n) > 0, \quad n \in \mathbb{Z}_{n_0-k}^{n_0-1}. \tag{4.1}$$

Hence, there is a constant $\gamma > 0$ such that

$$\Delta Y(n) - \gamma \Delta y(n) > 0, \quad n \in \mathbb{Z}_{n_0-k}^{n_0-1} \tag{4.2}$$

or

$$\Delta\big(Y(n) - \gamma y(n)\big) > 0, \quad n \in \mathbb{Z}_{n_0-k}^{n_0-1}, \tag{4.3}$$

and the function $Y(n) - \gamma y(n)$ is strictly increasing on $\mathbb{Z}_{n_0-k}^{n_0-1}$. Then Lemma 2.1 implies that $Y(n) - \gamma y(n)$ is strictly increasing on $\mathbb{Z}_{n_0-k}^{\infty}$. Thus

$$Y(n) - \gamma y(n) > Y(n_0) - \gamma y(n_0), \quad n \in \mathbb{Z}_{n_0}^{\infty} \tag{4.4}$$

or

$$y(n) < \frac{1}{\gamma}(Y(n) - Y(n_0)) + y(n_0), \quad n \in \mathbb{Z}_{n_0}^{\infty}, \tag{4.5}$$

and, consequently, $y(n)$ is a bounded function on $\mathbb{Z}_{n_0-k}^{\infty}$ because of the boundedness of $Y(n)$. Obviously, in such a case, $y(n)$ is asymptotically convergent and has a finite limit.

Summarizing the previous section, we state that every monotone solution is convergent. It remains to consider a class of all nonmonotone initial functions. For the behavior of a solution $y_{(n_0,\varphi)}$ generated by a nonmonotone initial function $\varphi \in \mathcal{C}$, there are two possibilities: $y_{(n_0,\varphi)}$ is either eventually monotone and, consequently, convergent, or $y_{(n_0,\varphi)}$ is eventually nonmonotone.

Now we use the statement of Lemma 2.5 that every discrete function $\varphi \in \mathcal{C}$ can be decomposed into the difference of two strictly increasing discrete functions $\varphi_j \in \mathcal{C}$, $j = 1,2$. In accordance with the previous part of the proof, every function $\varphi_j \in \mathcal{C}$, $j = 1,2$ defines a strictly increasing and asymptotically convergent solution $y_{(n_0,\varphi_j)}$. Now it is clear that the solution $y_{(n_0,\varphi)}$ is asymptotically convergent. □

We will finish the paper with two obvious results. Inequality (3.4) in Theorem 3.3 was necessary only for the proof of the existence of an asymptotically convergent solution. If we assume the existence of an asymptotically convergent solution rather than inequality (3.4), we can formulate the following result, the proof of which is an elementary modification of the proof of Theorem 4.1.

Theorem 4.2. *If* (1.1) *has a strictly monotone and asymptotically convergent solution on* $\mathbb{Z}_{n_0-k}^{\infty}$*, then all the solutions of* (1.1) *defined on* $\mathbb{Z}_{n_0-k}^{\infty}$ *are asymptotically convergent.*

to Contents, page 1

Combining the statements of Theorems 2.2, 3.1, and 4.2, we get a series of equivalent statements below.

Theorem 4.3. *The following three statements are equivalent.*

(a) *Equation* (1.1) *has a strictly monotone and asymptotically convergent solution on* $\mathbb{Z}^{\infty}_{n_0-k}$.

(b) *All solutions of* (1.1) *defined on* $\mathbb{Z}^{\infty}_{n_0-k}$ *are asymptotically convergent.*

(c) *Inequality* (2.1) *has a strictly monotone and asymptotically convergent solution on* $\mathbb{Z}^{\infty}_{n_0-k}$.

Example 4.4. We will demonstrate the sharpness of the criterion (3.4) by the following example. Let $k = 1$, $j = 0$, $\beta(n) = 1 - 1/n$, $n \in \mathbb{Z}^{\infty}_{n_0-1}$, $n_0 = 2$ in (1.1); that is, we consider the equation

$$\Delta y(n) = \left(1 - \frac{1}{n}\right)\left[y(n) - y(n-1)\right]. \tag{4.6}$$

By Theorems 3.3 and 4.3, all solutions are asymptotically convergent if

$$\beta(n) \le \frac{1}{k-j} - \frac{p(k+j+1)}{2n(k-j)} = 1 - \frac{p}{n}, \tag{4.7}$$

where a constant $p > 1$. In our case the inequality (4.7) does not hold since inequality

$$\beta(n) = 1 - \frac{1}{n} \le 1 - \frac{p}{n} \tag{4.8}$$

is valid only for $p \le 1$. Inequality (4.7) is sharp because there exists a solution $y = y^*(n)$ of (4.6) having the form of an nth partial sum of harmonic series, that is,

$$y^*(n) = \sum_{i=1}^{n} \frac{1}{i} \tag{4.9}$$

with the obvious property $\lim_{n\to\infty} y^*(n) = +\infty$. Then (by Theorem 4.3), all strictly monotone (increasing or decreasing) solutions of (4.6) tend to infinity.

Acknowledgments

The research was supported by the Project APVV-0700-07 of the Slovak Research and Development Agency and by the Grant no. 1/0090/09 of the Grant Agency of Slovak Republic (VEGA).

References

[1] O. Arino and M. Pituk, "Convergence in asymptotically autonomous functional-differential equations," *Journal of Mathematical Analysis and Applications*, vol. 237, no. 1, pp. 376–392, 1999.

to Contents, page 1

[2] H. Bereketoğlu and F. Karakoç, "Asymptotic constancy for impulsive delay differential equations," *Dynamic Systems and Applications*, vol. 17, no. 1, pp. 71–83, 2008.
[3] I. Györi, F. Karakoç, and H. Bereketoğlu, "Convergence of solutions of a linear impulsive differential equations system with many delays," *Dynamics of Continuous, Discrete and Impulsive Systems Series A, Mathematical Analysis*, vol. 18, no. 2, pp. 191–202, 2011.
[4] H. Bereketoğlu and M. Pituk, "Asymptotic constancy for nonhomogeneous linear differential equations with unbounded delays," *Discrete and Continuous Dynamical Systems. Series A*, supplement, pp. 100–107, 2003.
[5] J. Diblík, "Asymptotic convergence criteria of solutions of delayed functional differential equations," *Journal of Mathematical Analysis and Applications*, vol. 274, no. 1, pp. 349–373, 2002.
[6] J. Diblík and M. Růžičková, "Convergence of the solutions of the equation $\dot{y}(t) = \beta(t)[y(t-\delta)-y(t-\tau)]$ in the critical case," *The Journal of Mathematical Analysis and Applications*, vol. 331, pp. 1361–1370, 2007.
[7] E. Messina, Y. Muroya, E. Russo, and A. Vecchio, "Convergence of solutions for two delays Volterra integral equations in the critical case," *Applied Mathematics Letters*, vol. 23, no. 10, pp. 1162–1165, 2010.
[8] L. Berezansky and E. Braverman, "On oscillation of a food-limited population model with time delay," *Abstract and Applied Analysis*, no. 1, pp. 55–66, 2003.
[9] H. Bereketoglu and A. Huseynov, "Convergence of solutions of nonhomogeneous linear difference systems with delays," *Acta Applicandae Mathematicae*, vol. 110, no. 1, pp. 259–269, 2010.
[10] M. Dehghan and M. J. Douraki, "Global attractivity and convergence of a difference equation," *Dynamics of Continuous, Discrete and Impulsive Systems. Series A. Mathematical Analysis*, vol. 16, no. 3, pp. 347–361, 2009.
[11] I. Györi and L. Horváth, "Asymptotic constancy in linear difference equations: limit formulae and sharp conditions," *Advances in Difference Equations*, vol. 2010, Article ID 789302, 20 pages, 2010.
[12] S. Stević, "Global stability and asymptotics of some classes of rational difference equations," *Journal of Mathematical Analysis and Applications*, vol. 316, no. 1, pp. 60–68, 2006.
[13] C. M. Kent, W. Kosmala, and S. Stević, "On the difference equation $x_{n+1} = x_n x_{n-2} - 1$," *Abstract and Applied Analysis*, vol. 2011, Article ID 815285, 25 pages, 2011.
[14] J. Baštinec and J. Diblík, "Subdominant positive solutions of the discrete equation $\Delta u(k+n) = -p(k)u(k)$," *Abstract and Applied Analysis*, no. 6, pp. 461–470, 2004.
[15] R. Medina and M. Pituk, "Nonoscillatory solutions of a second-order difference equation of Poincaré type," *Applied Mathematics Letters*, vol. 22, no. 5, pp. 679–683, 2009.
[16] B. Z. Vulikh, *A Brief Course in the Theory of Functions of a Real Variable (An Introduction to the Theory of the Integral)*, Mir Publishers, Moscow, Russia, 1976.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 292860, 15 pages
doi:10.1155/2011/292860

Research Article

Boundary Value Problems for q-Difference Inclusions

Bashir Ahmad[1] and Sotiris K. Ntouyas[2]

[1] *Department of Mathematics, Faculty of Science, King Abdulaziz University, P.O. Box 80203, Jeddah 21589, Saudi Arabia*
[2] *Department of Mathematics, University of Ioannina, 451 10 Ioannina, Greece*

Correspondence should be addressed to Bashir Ahmad, bashir_qau@yahoo.com

Received 12 October 2010; Revised 24 January 2011; Accepted 22 February 2011

Academic Editor: Yuri V. Rogovchenko

We investigate the existence of solutions for a class of second-order q-difference inclusions with nonseparated boundary conditions. By using suitable fixed-point theorems, we study the cases when the right-hand side of the inclusions has convex as well as nonconvex values.

1. Introduction

The discretization of the ordinary differential equations is an important and necessary step towards finding their numerical solutions. Instead of the standard discretization based on the arithmetic progression, one can use an equally efficient q-discretization related to geometric progression. This alternative method leads to q-difference equations, which in the limit $q \to 1$ correspond to the classical differential equations. q-difference equations are found to be quite useful in the theory of quantum groups [1]. For historical notes and development of the subject, we refer the reader to [2, 3] while some recent results on q-difference equations can be found in [4–6]. However, the theory of boundary value problems for nonlinear q-difference equations is still in the initial stages, and many aspects of this theory need to be explored.

Differential inclusions arise in the mathematical modelling of certain problems in economics, optimal control, stochastic analysis, and so forth and are widely studied by many authors; see [7–13] and the references therein. For some works concerning difference inclusions and dynamic inclusions on time scales, we refer the reader to the papers [14–17].

to Contents, page 1

In this paper, we study the existence of solutions for second-order q-difference inclusions with nonseparated boundary conditions given by

$$D_q^2 u(t) \in F(t, u(t)), \quad 0 \leq t \leq T, \tag{1.1}$$

$$u(0) = \eta u(T), \qquad D_q u(0) = \eta D_q u(T), \tag{1.2}$$

where $F : [0,T] \times \mathbb{R} \to \mathcal{P}(\mathbb{R})$ is a compact valued multivalued map, $\mathcal{P}(\mathbb{R})$ is the family of all subsets of $\mathbb{R}$, T is a fixed constant, and $\eta \neq 1$ is a fixed real number.

The aim of our paper is to establish some existence results for the Problems (1.1)-(1.2), when the right-hand side is convex as well as nonconvex valued. First of all, an integral operator is found by applying the tools of q-difference calculus, which plays a pivotal role to convert the given boundary value problem to a fixed-point problem. Our approach is simpler as it does not involve the typical series solution form of q-difference equations. The first result relies on the nonlinear alternative of Leray-Schauder type. In the second result, we will combine the nonlinear alternative of Leray-Schauder type for single-valued maps with a selection theorem due to Bressan and Colombo for lower semicontinuous multivalued maps with nonempty closed and decomposable values, while in the third result, we will use the fixed-point theorem for generalized contraction multivalued maps due to Wegrzyk. The methods used are standard; however, their exposition in the framework of Problems (1.1)-(1.2) is new.

The paper is organized as follows: in Section 2, we recall some preliminary facts that we need in the sequel, and we prove our main results in Section 3.

2. Preliminaries

In this section, we introduce notation, definitions, and preliminary facts which we need for the forthcoming analysis.

2.1. q-Calculus

Let us recall some basic concepts of q-calculus [1–3].

For $0 < q < 1$, we define the q-derivative of a real-valued function f as

$$D_q f(t) = \frac{f(t) - f(qt)}{(1-q)t}, \qquad D_q f(0) = \lim_{t \to 0} D_q f(t). \tag{2.1}$$

The higher-order q-derivatives are given by

$$D_q^0 f(t) = f(t), \qquad D_q^n f(t) = D_q D_q^{n-1} f(t), \quad n \in \mathbb{N}. \tag{2.2}$$

The q-integral of a function f defined in the interval $[a,b]$ is given by

$$\int_a^x f(t) d_q t := \sum_{n=0}^{\infty} x(1-q)q^n f(xq^n) - af(q^n a), \quad x \in [a,b], \tag{2.3}$$

to Contents, page 1

and for $a = 0$, we denote

$$I_q f(x) = \int_0^x f(t) d_q t = \sum_{n=0}^{\infty} x(1-q) q^n f(xq^n), \tag{2.4}$$

provided the series converges. If $a \in [0, b]$ and f is defined in the interval $[0, b]$, then

$$\int_a^b f(t) d_q t = \int_0^b f(t) d_q t - \int_0^a f(t) d_q t. \tag{2.5}$$

Similarly, we have

$$I_q^0 f(t) = f(t), \qquad I_q^n f(t) = I_q I_q^{n-1} f(t), \quad n \in \mathbb{N}. \tag{2.6}$$

Observe that

$$D_q I_q f(x) = f(x), \tag{2.7}$$

and if f is continuous at $x = 0$, then

$$I_q D_q f(x) = f(x) - f(0). \tag{2.8}$$

In q-calculus, the integration by parts formula is

$$\int_0^x f(t) D_q g(t) d_q t = \left[f(t) g(t) \right]_0^x - \int_0^x D_q f(t) g(qt) d_q t. \tag{2.9}$$

2.2. Multivalued Analysis

Let us recall some basic definitions on multivalued maps [18, 19].

Let X denote a normed space with the norm $|\cdot|$. A multivalued map $G : X \to \mathcal{P}(X)$ is convex (closed) valued if $G(x)$ is convex (closed) for all $x \in X$. G is bounded on bounded sets if $G(B) = \cup_{x \in B} G(x)$ is bounded in X for all bounded sets B in X (i.e., $\sup_{x \in B}\{\sup\{|y| : y \in G(x)\}\} < \infty$). G is called upper semicontinuous (u.s.c.) on X if for each $x_0 \in X$, the set $G(x_0)$ is a nonempty closed subset of X, and if for each open set N of X containing $G(x_0)$, there exists an open neighborhood N_0 of x_0 such that $G(N_0) \subseteq N$. G is said to be completely continuous if $G(\mathcal{B})$ is relatively compact for every bounded set B in X. If the multivalued map G is completely continuous with nonempty compact values, then G is u.s.c. if and only if G has a closed graph (i.e., $x_n \to x_*$, $y_n \to y_*$, $y_n \in G(x_n)$ imply $y_* \in G(x_*)$). G has a fixed-point if there is $x \in X$ such that $x \in G(x)$. The fixed-point set of the multivalued operator G will be denoted by $\operatorname{Fix} G$.

For more details on multivalued maps, see the books of Aubin and Cellina [20], Aubin and Frankowska [21], Deimling [18], and Hu and Papageorgiou [19].

to Contents, page 1

Let $C([0,T],\mathbb{R})$ denote the Banach space of all continuous functions from $[0,T]$ into $\mathbb{R}$ with the norm

$$\|u\|_\infty = \sup\{|u(t)| : t \in [0,T]\}. \tag{2.10}$$

Let $L^1([0,T],\mathbb{R})$ be the Banach space of measurable functions $u : [0,T] \to \mathbb{R}$ which are Lebesgue integrable and normed by

$$\|u\|_{L^1} = \int_0^T |u(t)|dt, \quad \forall u \in L^1([0,T],\mathbb{R}). \tag{2.11}$$

Definition 2.1. A multivalued map $G : [0,T] \to \mathcal{P}(\mathbb{R})$ with nonempty compact convex values is said to be measurable if for any $x \in \mathbb{R}$, the function

$$\mathrm{t} \longmapsto d(x, F(t)) = \inf\{|x - z| : z \in F(t)\} \tag{2.12}$$

is measurable.

Definition 2.2. A multivalued map $F : [0,T] \times \mathbb{R} \to \mathcal{P}(\mathbb{R})$ is said to be Carathéodory if

(i) $t \mapsto F(t,x)$ is measurable for each $x \in \mathbb{R}$,

(ii) $x \mapsto F(t,x)$ is upper semicontinuous for almost all $t \in [0,T]$.

Further a Carathéodory function F is called L^1-Carathéodory if

(iii) for each $\alpha > 0$, there exists $\varphi_\alpha \in L^1([0,T],\mathbb{R}^+)$ such that

$$\|F(t,x)\| = \sup\{|v| : v \in F(t,x)\} \le \varphi_\alpha(t) \tag{2.13}$$

for all $\|x\|_\infty \le \alpha$ and for a.e. $t \in [0,T]$.

Let E be a Banach space, let X be a nonempty closed subset of E, and let $G : X \to \mathcal{P}(E)$ be a multivalued operator with nonempty closed values. G is lower semicontinuous (l.s.c.) if the set $\{x \in X : G(x) \cap B \neq \emptyset\}$ is open for any open set B in E. Let A be a subset of $[0,T] \times \mathbb{R}$. A is $\mathcal{L} \otimes \mathcal{B}$ measurable if A belongs to the σ-algebra generated by all sets of the form $\mathcal{J} \times D$, where $\mathcal{J}$ is Lebesgue measurable in $[0,T]$ and D is Borel measurable in $\mathbb{R}$. A subset A of $L^1([0,T],\mathbb{R})$ is decomposable if for all $u, v \in A$ and $\mathcal{J} \subset [0,T]$ measurable, the function $u\chi_{\mathcal{J}} + v\chi_{J-\mathcal{J}} \in A$, where $\chi_{\mathcal{J}}$ stands for the characteristic function of $\mathcal{J}$.

Definition 2.3. If $F : [0,T] \times \mathbb{R} \to \mathcal{P}(\mathbb{R})$ is a multivalued map with compact values and $u(\cdot) \in C([0,\mathrm{T}],\mathbb{R})$, then $F(\cdot,\cdot)$ is of lower semicontinuous type if

$$S_F(u) = \left\{ w \in L^1([0,T],\mathbb{R}) : w(t) \in F(t,u(t)) \quad \text{for a.e. } t \in [0,T] \right\} \tag{2.14}$$

is lower semicontinuous with closed and decomposable values.

to Contents, page 1

Let (X, d) be a metric space associated with the norm $|\cdot|$. The Pompeiu-Hausdorff distance of the closed subsets $A, B \subset X$ is defined by

$$d_H(A, B) = \max\{d^*(A, B), d^*(B, A)\}, d^*(A, B) = \sup\{d(a, B) : a \in A\}, \tag{2.15}$$

where $d(x, B) = \inf_{y \in B} d(x, y)$.

Definition 2.4. A function $l : \mathbb{R}_+ \to \mathbb{R}_+$ is said to be a strict comparison function (see [25]) if it is continuous strictly increasing and $\sum_{n=1}^{\infty} l^n(t) < \infty$, for each $t > 0$.

Definition 2.5. A multivalued operator N on X with nonempty values in X is called

(a) γ-Lipschitz if and only if there exists $\gamma > 0$ such that

$$d_H(N(x), N(y)) \le \gamma d(x, y), \quad \text{for each } x, y \in X, \tag{2.16}$$

(b) a contraction if and only if it is γ-Lipschitz with $\gamma < 1$,

(c) a generalized contraction if and only if there is a strict comparison function $l : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$d_H(N(x), N(y)) \le l(d(x, y)), \quad \text{for each } x, y \in X. \tag{2.17}$$

The following lemmas will be used in the sequel.

Lemma 2.6 (see [22]). *Let X be a Banach space. Let $F : [0, T] \times X \to \mathcal{P}(X)$ be an L^1-Carathéodory multivalued map with $S_F \ne \emptyset$, and let Γ be a linear continuous mapping from $L^1([0, T], X)$ to $C([0, T], X)$, then the operator*

$$\Gamma \circ S_F : C([0, T], X) \longrightarrow \mathcal{P}(C([0, T], X)) \tag{2.18}$$

defined by $(\Gamma \circ S_F)(x) = \Gamma(S_F(x))$ has compact convex values and has a closed graph operator in $C([0, T], X) \times C([0, T], X)$.

In passing, we remark that if $\dim X < \infty$, then $S_F(x) \ne \emptyset$ for any $x(\cdot) \in C([0, T], X)$ with $F(\cdot, \cdot)$ as in Lemma 2.6.

Lemma 2.7 (nonlinear alternative for Kakutani maps [23]). *Let E be a Banach space, C, a closed convex subset of E, U an open subset of C and $0 \in U$. Suppose that $F : \overline{U} \to \mathcal{P}_{c,cv}(C)$ is an upper semicontinuous compact map; here, $\mathcal{P}_{c,cv}(C)$ denotes the family of nonempty, compact convex subsets of C, then either*

(i) *F has a fixed-point in $\overline{U}$,*

(ii) *or there is a $u \in \partial U$ and $\lambda \in (0, 1)$ with $u \in \lambda F(u)$.*

Lemma 2.8 (see [24]). *Let Y be a separable metric space, and let $N : Y \to \mathcal{P}(L^1([0, T], \mathbb{R}))$ be a lower semicontinuous multivalued map with closed decomposable values, then $N(\cdot)$ has a continuous*

to Contents, page 1

selection; that is, there exists a continuous mapping (single-valued) $g : Y \to L^1([0,T],\mathbb{R})$ *such that* $g(y) \in N(y)$ *for every* $y \in Y$.

Lemma 2.9 (Wegrzyk's fixed-point theorem [25, 26]). *Let* (X,d) *be a complete metric space. If* $N : X \to \mathcal{P}(X)$ *is a generalized contraction with nonempty closed values, then* $\mathrm{Fix}N \neq \emptyset$.

Lemma 2.10 (Covitz and Nadler's fixed-point theorem [27]). *Let* (X,d) *be a complete metric space. If* $N : X \to \mathcal{P}(X)$ *is a multivalued contraction with nonempty closed values, then* N *has a fixed-point* $z \in X$ *such that* $z \in N(z)$*, that is,* $\mathrm{Fix}N \neq \emptyset$.

3. Main Results

In this section, we are concerned with the existence of solutions for the Problems (1.1)-(1.2) when the right-hand side has convex as well as nonconvex values. Initially, we assume that F is a compact and convex valued multivalued map.

To define the solution for the Problems (1.1)-(1.2), we need the following result.

Lemma 3.1. *Suppose that* $\sigma : [0,T] \to \mathbb{R}$ *is continuous, then the following problem*

$$\begin{gathered} D_q^2 u(t) = \sigma(t), \quad \text{a.e. } t \in [0,T], \\ u(0) = \eta u(T), \qquad D_q u(0) = \eta D_q u(T) \end{gathered} \tag{3.1}$$

has a unique solution

$$u(t) = \int_0^T G(t,qs)\sigma(s)d_q s, \tag{3.2}$$

where $G(t,qs)$ *is the Green's function given by*

$$G(t,qs) = \frac{1}{(\eta-1)^2}\begin{cases} \eta(\eta-1)(qs-t) + \eta T, & \text{if } 0 \le t < s \le T, \\ (\eta-1)(qs-t) + \eta T, & \text{if } 0 \le s \le t \le T. \end{cases} \tag{3.3}$$

Proof. In view of (2.7) and (2.9), the solution of $D_q^2 u = \sigma(t)$ can be written as

$$u(t) = \int_0^t (t-qs)\sigma(s)d_q s + a_1 t + a_2, \tag{3.4}$$

to Contents, page 1

where a_1, a_2 are arbitrary constants. Using the boundary conditions (1.2) and (3.4), we find that

$$a_1 = \frac{-\eta}{(\eta-1)} \int_0^T \sigma(s) d_q s,$$
$$a_2 = \frac{\eta^2 T}{(\eta-1)^2} \int_0^T \sigma(s) d_q s - \frac{\eta}{(\eta-1)} \int_0^T (T-qs)\sigma(s) d_q s. \tag{3.5}$$

Substituting the values of a_1 and a_2 in (3.4), we obtain (3.2). □

Let us denote

$$G_1 = \max_{t,s\in[0,T]} |G(t,qs)|. \tag{3.6}$$

Definition 3.2. A function $u \in C([0,T],\mathbb{R})$ is said to be a solution of (1.1)-(1.2) if there exists a function $v \in L^1([0,T],\mathbb{R})$ with $v(t) \in F(t,x(t))$ a.e. $t \in [0,T]$ and

$$u(t) = \int_0^T G(t,qs)v(s) d_q s, \tag{3.7}$$

where $G(t,qs)$ is given by (3.3).

Theorem 3.3. *Suppose that*

(H1) *the map* $F : [0,T]\times\mathbb{R} \to \mathcal{P}(\mathbb{R})$ *has nonempty compact convex values and is Carathéodory,*

(H2) *there exist a continuous nondecreasing function* $\psi : [0,\infty) \to (0,\infty)$ *and a function* $p \in L^1([0,T],\mathbb{R}_+)$ *such that*

$$\|F(t,u)\|_{\mathcal{P}} := \sup\{|v| : v \in F(t,u)\} \le p(t)\psi(\|u\|_\infty) \tag{3.8}$$

for each $(t,u) \in [0,T]\times\mathbb{R}$,

(H3) *there exists a number* $M > 0$ *such that*

$$\frac{M}{G_1\psi(M)\|p\|_{L^1}} > 1, \tag{3.9}$$

then the BVP (1.1)-(1.2) has at least one solution.

Proof. In view of Definition 3.2, the existence of solutions to (1.1)-(1.2) is equivalent to the existence of solutions to the integral inclusion

$$u(t) \in \int_0^T G(t,qs)F(s,u(s)) d_q s, \quad t \in [0,T]. \tag{3.10}$$

to Contents, page 1

Let us introduce the operator

$$N(u) := \left\{ h \in C([0,T],\mathbb{R}) : h(t) = \int_0^T G(t,qs)v(s)d_qs,\ v \in S_{F,u} \right\}. \tag{3.11}$$

We will show that N satisfies the assumptions of the nonlinear alternative of Leray-Schauder type. The proof will be given in several steps.

Step 1 ($N(u)$ is convex for each $u \in C([0,T],\mathbb{R})$). Indeed, if h_1, h_2 belong to $N(u)$, then there exist $v_1, v_2 \in S_{F,u}$ such that for each $t \in [0,T]$, we have

$$h_i(t) = \int_0^T G(t,qs)v_i(s)d_qs, \quad (i = 1,2). \tag{3.12}$$

Let $0 \le d \le 1$, then, for each $t \in [0,T]$, we have

$$(dh_1 + (1-d)h_2)(t) = \int_0^T G(t,qs)[dv_1(s) + (1-d)v_2(s)]d_qs. \tag{3.13}$$

Since $S_{F,u}$ is convex (because F has convex values); therefore,

$$dh_1 + (1-d)h_2 \in N(u). \tag{3.14}$$

Step 2 (N maps bounded sets into bounded sets in $C([0,T],\mathbb{R})$). Let $B_m = \{u \in C([0,T],\mathbb{R}) : \|u\|_\infty \le m, m > 0\}$ be a bounded set in $C([0,T],\mathbb{R})$ and $u \in B_m$, then for each $h \in N(u)$, there exists $v \in S_{F,u}$ such that

$$h(t) = \int_0^T G(t,qs)v(s)d_qs. \tag{3.15}$$

Then, in view of (H2), we have

$$\begin{aligned} |h(t)| &\le \int_0^T \left|G(t,qs)\right||v(s)|d_qs \\ &\le G_1 \int_0^T p(s)\psi(\|u\|_\infty)d_qs \\ &\le G_1\psi(m)\int_0^T p(s)d_qs. \end{aligned} \tag{3.16}$$

Thus,

$$\|h\|_\infty \le G_1\psi(m)\left\|p\right\|_{L^1}. \tag{3.17}$$

to Contents, page 1

Step 3 (N maps bounded sets into equicontinuous sets of $C([0,T],\mathbb{R})$). Let $r_1, r_2 \in [0, T]$, $r_1 < r_2$ and B_m be a bounded set of $C([0,T],\mathbb{R})$ as in Step 2 and $x \in B_m$. For each $h \in N(u)$

$$\begin{aligned}
|h(r_2) - h(r_1)| &\leq \int_0^T |G(r_2,s) - G(r_1,s)||v(s)|d_q s \\
&\leq \varphi(\|u\|_\infty) \int_0^T |G(r_2,s) - G(r_1,s)|p(s)d_q s \\
&\leq \varphi(m) \int_0^T |G(r_2,s) - G(r_1,s)|p(s)d_q s.
\end{aligned} \tag{3.18}$$

The right-hand side tends to zero as $r_2 - r_1 \to 0$. As a consequence of Steps 1 to 3 together with the Arzelá-Ascoli Theorem, we can conclude that $N : C([0,T],\mathbb{R}) \to \mathcal{P}(C([0,T],\mathbb{R}))$ is completely continuous.

Step 4 (N has a closed graph). Let $u_n \to u_*$, $h_n \in N(u_n)$, and $h_n \to h_*$. We need to show that $h_* \in N(u_*)$. $h_n \in N(u_n)$ means that there exists $v_n \in S_{F,u_n}$ such that, for each $t \in [0,T]$,

$$h_n(t) = \int_0^T G(t,qs)v_n(s)d_q s. \tag{3.19}$$

We must show that there exists $h_* \in S_{F,u_*}$ such that, for each $t \in [0,T]$,

$$h_*(t) = \int_0^T G(t,qs)v_*(s)d_q s. \tag{3.20}$$

Clearly, we have

$$\|h_n - h_*\|_\infty \longrightarrow 0 \quad \text{as } n \longrightarrow \infty. \tag{3.21}$$

Consider the continuous linear operator

$$\Gamma : L^1([0,T],\mathbb{R}) \longrightarrow C([0,T],\mathbb{R}), \tag{3.22}$$

defined by

$$v \longmapsto (\Gamma v)(t) = \int_0^T G(t,qs)v(s)d_q s. \tag{3.23}$$

From Lemma 2.6, it follows that $\Gamma \circ S_F$ is a closed graph operator. Moreover, we have

$$h_n(t) \in \Gamma(S_{F,u_n}). \tag{3.24}$$

to Contents, page 1

Since $u_n \to u_*$, it follows from Lemma 2.6 that

$$h_*(t) = \int_0^T G(t, qs)v_*(s)d_q s \tag{3.25}$$

for some $v_* \in S_{F,u_*}$.

Step 5 (a priori bounds on solutions). Let u be a possible solution of the Problems (1.1)-(1.2), then there exists $v \in L^1([0,T],\mathbb{R})$ with $v \in S_{F,u}$ such that, for each $t \in [0,T]$,

$$u(t) = \int_0^T G(t, qs)v(s)d_q s. \tag{3.26}$$

For each $t \in [0,T]$, it follows by (H2) and (H3) that

$$\begin{aligned} |u(t)| &\le G_1 \int_0^T p(s)\psi(\|u\|_\infty)d_q s \\ &\le G_1\psi(\|u\|_\infty)\int_0^T p(s)d_q s. \end{aligned} \tag{3.27}$$

Consequently,

$$\frac{\|u\|_\infty}{G_1\psi(\|u\|_\infty)\|p\|_{L^1}} \le 1. \tag{3.28}$$

Then by (H3), there exists M such that $\|u\|_\infty \ne M$.

Let

$$U = \{u \in C([0,T],\mathbb{R}) : \|u\|_\infty < M+1\}. \tag{3.29}$$

The operator $N : \overline{U} \to \mathcal{P}(C([0,T],\mathbb{R}))$ is upper semicontinuous and completely continuous. From the choice of U, there is no $u \in \partial U$ such that $u \in \lambda N(u)$ for some $\lambda \in (0,1)$. Consequently, by Lemma 2.7, it follows that N has a fixed-point u in $\overline{U}$ which is a solution of the Problems (1.1)-(1.2). This completes the proof. □

Next, we study the case where F is not necessarily convex valued. Our approach here is based on the nonlinear alternative of Leray-Schauder type combined with the selection theorem of Bressan and Colombo for lower semicontinuous maps with decomposable values.

Theorem 3.4. *Suppose that the conditions (H2) and (H3) hold. Furthermore, it is assumed that*

(H4) $F : [0,T]\times\mathbb{R} \to \mathcal{P}(\mathbb{R})$ *has nonempty compact values and*

(a) $(t,u) \mapsto F(t,u)$ *is* $\mathcal{L}\otimes\mathcal{B}$ *measurable,*
(b) $u \mapsto F(t,u)$ *is lower semicontinuous for a.e.* $t \in [0,T]$*,*

to Contents, page 1

(H5) *for each* $\rho > 0$, *there exists* $\varphi_\rho \in L^1([0,T],\mathbb{R}_+)$ *such that*

$$\|F(t,u)\| = \sup\{|v| : v \in F(t,u)\} \le \varphi_\rho(t) \quad \forall \|u\|_\infty \le \rho \text{ and for a.e. } t \in [0,T]. \tag{3.30}$$

then, the BVP (1.1)-(1.2) *has at least one solution.*

Proof. Note that (H4) and (H5) imply that F is of lower semicontinuous type. Thus, by Lemma 2.8, there exists a continuous function $f : C([0,T],\mathbb{R}) \to L^1([0,T],\mathbb{R})$ such that $f(u) \in \mathcal{F}(u)$ for all $u \in C([0,T],\mathbb{R})$. So we consider the problem

$$\begin{gathered} D_q^2 u(t) = f(u(t)), \quad 0 \le t \le T, \\ u(0) = \eta u(T), \qquad D_q u(0) = \eta D_q u(T). \end{gathered} \tag{3.31}$$

Clearly, if $u \in C([0,T],\mathbb{R})$ is a solution of (3.31), then u is a solution to the Problems (1.1)-(1.2). Transform the Problem (3.31) into a fixed-point theorem

$$u(t) = \left(\overline{N}u\right)(t), \quad t \in [0,T], \tag{3.32}$$

where

$$\left(\overline{N}u\right)(t) = \int_0^T G(t,qs) f(u(s)) d_q s, \quad t \in [0,T]. \tag{3.33}$$

We can easily show that $\overline{N}$ is continuous and completely continuous. The remainder of the proof is similar to that of Theorem 3.3. □

Now, we prove the existence of solutions for the Problems (1.1)-(1.2) with a nonconvex valued right-hand side by applying Lemma 2.9 due to Wegrzyk.

Theorem 3.5. *Suppose that*

(H6) $F : [0,T] \times \mathbb{R} \to \mathcal{P}(\mathbb{R})$ *has nonempty compact values and* $F(\cdot,u)$ *is measurable for each* $u \in \mathbb{R}$,

(H7) $d_H(F(t,u),F(t,\overline{u})) \le k(t) l(|u-\overline{u}|)$ *for almost all* $t \in [0,1]$ *and* $u,\overline{u} \in \mathbb{R}$ *with* $k \in L^1([0,1],\mathbb{R}_+)$ *and* $d(0,F(t,0)) \le k(t)$ *for almost all* $t \in [0,1]$, *where* $l : \mathbb{R}_+ \to \mathbb{R}_+$ *is strictly increasing,*

then the BVP (1.1)-(1.2) *has at least one solution on* $[0,T]$ *if* $\gamma l : \mathbb{R}_+ \to \mathbb{R}_+$ *is a strict comparison function, where* $\gamma = G_1 \|k\|_{L^1}$.

to Contents, page 1

Proof. Suppose that $\gamma l : \mathbb{R}_+ \to \mathbb{R}_+$ is a strict comparison function. Observe that by the assumptions (H6) and (H7), $F(\cdot, u(\cdot))$ is measurable and has a measurable selection $v(\cdot)$ (see Theorem 3.6 [28]). Also $k \in L^1([0,1],\mathbb{R})$ and

$$\begin{aligned}|v(t)| &\le d(0,F(t,0)) + H_d(F(t,0),F(t,u(t))) \\ &\le k(t) + k(t)l(|u(t)|) \\ &\le (1 + l(\|u\|_\infty))k(t).\end{aligned} \tag{3.34}$$

Thus, the set $S_{F,u}$ is nonempty for each $u \in C([0,T],\mathbb{R})$.

As before, we transform the Problems (1.1)-(1.2) into a fixed-point problem by using the multivalued map N given by (3.11) and show that the map N satisfies the assumptions of Lemma 2.9. To show that the map $N(u)$ is closed for each $u \in C([0,T],\mathbb{R})$, let $(u_n)_{n\ge 0} \in N(u)$ such that $u_n \to \tilde{u}$ in $C([0,T],\mathbb{R})$, then $\tilde{u} \in C([0,T],\mathbb{R})$ and there exists $v_n \in S_{F,u}$ such that, for each $t \in [0,T]$,

$$u_n(t) = \int_0^T G(\mathrm{t}, qs)v_n(s)d_qs. \tag{3.35}$$

As F has compact values, we pass onto a subsequence to obtain that v_n converges to v in $L^1([0,T],\mathbb{R})$. Thus, $v \in S_{F,u}$ and for each $t \in [0,T]$,

$$u_n(t) \longrightarrow \tilde{u}(t) = \int_0^T G(t,qs)v(s)d_qs. \tag{3.36}$$

So, $\tilde{u} \in N(u)$ and hence $N(u)$ is closed.

Next, we show that

$$d_H(N(u),N(\overline{u})) \le \gamma l(\|u - \overline{u}\|_\infty) \quad \text{for each } u,\overline{u} \in C([0,T],\mathbb{R}). \tag{3.37}$$

Let $u,\overline{u} \in C([0,T],\mathbb{R})$ and $h_1 \in N(u)$. Then, there exists $v_1(t) \in S_{F,u}$ such that for each $t \in [0,T]$,

$$h_1(t) = \int_0^T G(t,qs)v_1(s)d_qs. \tag{3.38}$$

From (H7), it follows that

$$d_H(F(t,u(t)),F(t,\overline{u}(t))) \le k(t)l(|u(t) - \overline{u}(t)|). \tag{3.39}$$

So, there exists $w \in F(t,\overline{u}(t))$ such that

$$|v_1(t) - w| \le k(t)l(|u(t) - \overline{u}(t)|), \quad t \in [0,T]. \tag{3.40}$$

to Contents, page 1

Define $U : [0,T] \to \mathcal{P}(\mathbb{R})$ as

$$U(t) = \{w \in \mathbb{R} : |v_1(t) - w| \leq k(t)l(|u(t) - \overline{u}(t)|)\}. \tag{3.41}$$

Since the multivalued operator $U(t) \cap F(t, \overline{u}(t))$ is measurable (see Proposition 3.4 in [28]), there exists a function $v_2(t)$ which is a measurable selection for $U(t) \cap F(t, \overline{u}(t))$. So, $v_2(t) \in F(t, \overline{u}(t))$, and for each $t \in [0,T]$,

$$|v_1(t) - v_2(t)| \leq k(t)l(|u(t) - \overline{u}(t)|). \tag{3.42}$$

For each $t \in [0,T]$, let us define

$$h_2(t) = \int_0^T G(t, qs)v_2(s)d_q s, \tag{3.43}$$

then

$$\begin{aligned} |h_1(t) - h_2(t)| &\leq \int_0^T |G(t, qs)||v_1(s) - v_2(s)|d_q s \\ &\leq G_1 \int_0^T k(s)l(\|u - \overline{u}\|)d_q s. \end{aligned} \tag{3.44}$$

Thus,

$$\|h_1 - h_2\|_\infty \leq G_1 \|k\|_{L^1} l(\|u - \overline{u}\|_\infty) = \gamma l(\|u - \overline{u}\|_\infty). \tag{3.45}$$

By an analogous argument, interchanging the roles of u and $\overline{u}$, we obtain

$$d_H(N(u), N(\overline{u})) \leq G_1 \|k\|_{L^1} l(\|u - \overline{u}\|_\infty) = \gamma l(\|u - \overline{u}\|_\infty) \tag{3.46}$$

for each $u, \overline{u} \in C([0,T], \mathbb{R})$. So, N is a generalized contraction, and thus, by Lemma 2.9, N has a fixed-point u which is a solution to (1.1)-(1.2). This completes the proof. □

Remark 3.6. We notice that Theorem 3.5 holds for several values of the function l, for example, $l(t) = \ln(1+t)/\chi$, where $\chi \in (0,1)$, $l(t) = t$, and so forth. Here, we emphasize that the condition (H7) reduces to $d_H(F(t,u), F(t,\overline{u})) \leq k(t)|u - \overline{u}|$ for $l(t) = t$, where a contraction principle for multivalued map due to Covitz and Nadler [27] (Lemma 2.10) is applicable under the condition $G_1 \|k\|_{L^1} < 1$. Thus, our result dealing with a nonconvex valued right-hand side of (1.1) is more general, and the previous results for nonconvex valued right-hand side of the inclusions based on Covitz and Nadler's fixed-point result (e.g., see [8]) can be extended to generalized contraction case.

to Contents, page 1

Remark 3.7. Our results correspond to the ones for second-order q-difference inclusions with antiperiodic boundary conditions $(u(0) = -u(T), D_q u(0) = -D_q u(T))$ for $\eta = -1$. The results for an initial value problem of second-order q-difference inclusions follow for $\eta = 0$. These results are new in the present configuration.

Remark 3.8. In the limit $q \to 1$, the obtained results take the form of their "continuous" (i.e., differential) counterparts presented in Sections 4 (ii) for $\lambda_1 = \lambda_2 = \eta, \mu_1 = 0 = \mu_2$ of [29].

Example 3.9. Consider a boundary value problem of second-order q-difference inclusions given by

$$\begin{gathered} D_q^2 u(t) \in F(t, u(t)), \quad 0 \le t \le 1 \\ u(0) = -u(1), \qquad D_q u(0) = -D_q u(1), \end{gathered} \tag{3.47}$$

where $\eta = -1$ and $F : [0,1] \times \mathbb{R} \to \mathcal{P}(\mathbb{R})$ is a multivalued map given by

$$(t, u) \longrightarrow F(t, u) = \left[\frac{u^3}{u^3 + 3} + t^3 + 3, \frac{u}{u+1} + t + 1\right]. \tag{3.48}$$

For $f \in F$, we have

$$|f| \le \max\left(\frac{u^3}{u^3 + 3} + t^3 + 3, \frac{u}{u+1} + t + 1\right) \le 5, \quad u \in \mathbb{R}. \tag{3.49}$$

Thus,

$$\|F(t, u)\|_{\mathcal{P}} := \sup\{|y| : y \in F(t, u)\} \le 5 = p(t)\psi(\|u\|_\infty), \quad u \in \mathbb{R}, \tag{3.50}$$

with $p(t) = 1, \psi(\|u\|_\infty) = 5$. Further, using the condition

$$\frac{M}{G_1 \psi(M) \|p\|_{L^1}} > 1, \tag{3.51}$$

we find that $M > 5G_2$, where $G_2 = G_1|_{\eta=-1,T=1}$. Clearly, all the conditions of Theorem 3.3 are satisfied. So, the conclusion of Theorem 3.3 applies to the Problem (3.47).

Acknowledgments

The authors thank the referees for their comments. The research of B. Ahmad was partially supported by Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, Saudi Arabia.

to Contents, page 1

References

[1] V. Kac and P. Cheung, *Quantum Calculus*, Springer, New York, NY, USA, 2002.
[2] F. H. Jackson, "On q-definite integrals," *Quarterly Journal*, vol. 41, pp. 193–203, 1910.
[3] G. Gasper and M. Rahman, *Basic Hypergeometric Series*, Cambridge University Press, Cambridge, UK, 1990.
[4] M. H. Annaby and Z. S. Mansour, "q-Taylor and interpolation series for Jackson q-difference operators," *Journal of Mathematical Analysis and Applications*, vol. 344, no. 1, pp. 472–483, 2008.
[5] A. Dobrogowska and A. Odzijewicz, "Second order q-difference equations solvable by factorization method," *Journal of Computational and Applied Mathematics*, vol. 193, no. 1, pp. 319–346, 2006.
[6] M. El-Shahed and H. A. Hassan, "Positive solutions of q-Difference equation," *Proceedings of the American Mathematical Society*, vol. 138, no. 5, pp. 1733–1738, 2010.
[7] E. O. Ayoola, "Quantum stochastic differential inclusions satisfying a general lipschitz condition," *Dynamic Systems and Applications*, vol. 17, no. 3-4, pp. 487–502, 2008.
[8] A. Belarbi and M. Benchohra, "Existence results for nonlinear boundary-value problems with integral boundary conditions," *Electronic Journal of Differential Equations*, vol. 2005, no. 06, pp. 1–10, 2005.
[9] M. Benaïm, J. Hofbauer, and S. Sorin, "Stochastic approximations and differential inclusions, Part II: applications," *Mathematics of Operations Research*, vol. 31, no. 4, pp. 673–695, 2006.
[10] Y. K. Chang, W. T. Li, and J. J. Nieto, "Controllability of evolution differential inclusions in Banach spaces," *Nonlinear Analysis, Theory, Methods and Applications*, vol. 67, no. 2, pp. 623–632, 2007.
[11] S. K. Ntouyas, "Neumann boundary value problems for impulsive differential inclusions," *Electronic Journal of Qualitative Theory of Differential Equations*, no. 22, pp. 1–13, 2009.
[12] J. Simsen and C. B. Gentile, "Systems of p-Laplacian differential inclusions with large diffusion," *Journal of Mathematical Analysis and Applications*, vol. 368, no. 2, pp. 525–537, 2010.
[13] G. V. Smirnov, *Introduction to the Theory of Differential Inclusions*, American Mathematical Society, Providence, RI, USA, 2002.
[14] N. Apreutesei and G. Apreutesei, "A Trotter-Kato type result for a second order difference inclusion in a Hilbert space," *Journal of Mathematical Analysis and Applications*, vol. 361, no. 1, pp. 195–204, 2010.
[15] F. M. Atici and D. C. Biles, "First order dynamic inclusions on time scales," *Journal of Mathematical Analysis and Applications*, vol. 292, no. 1, pp. 222–237, 2004.
[16] A. Cernea and C. Georgescu, "Necessary optimality conditions for differential-difference inclusions with state constraints," *Journal of Mathematical Analysis and Applications*, vol. 334, no. 1, pp. 43–53, 2007.
[17] Y. K. Chang and W. T. Li, "Existence results for second-order dynamic inclusion with m-point boundary value conditions on time scales," *Applied Mathematics Letters*, vol. 20, no. 8, pp. 885–891, 2007.
[18] K. Deimling, *Multivalued Differential Equations*, Walter De Gruyter, Berlin, Germany, 1992.
[19] S. Hu and N. Papageorgiou, *Handbook of Multivalued Analysis, Theory I*, Kluwer Academic, Dodrecht, The Netherlands, 1997.
[20] J. P. Aubin and A. Cellina, *Differential Inclusions*, Springer, Heidelberg, Germany, 1984.
[21] J. P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäauser, Boston, Mass, USA, 1990.
[22] A. Lasota and Z. Opial, "An application of the Kakutani-Ky Fan theorem in the theory of ordinary differential equations," *Bulletin de l'Académie Polonaise des Sciences. Série des Sciences Mathématiques, Astronomiques et Physiques*, vol. 13, pp. 781–786, 1965.
[23] A. Granas and J. Dugundji, *Fixed Point Theory*, Springer, New York, NY, USA, 2003.
[24] A. Bressan and G. Colombo, "Extensions and selections of maps with decomposable values," *Studia Mathematica*, vol. 90, pp. 69–86, 1988.
[25] T. A. Lazăr, A. Petruşel, and N. Shahzad, "Fixed points for non-self operators and domain invariance theorems," *Nonlinear Analysis, Theory, Methods and Applications*, vol. 70, no. 1, pp. 117–125, 2009.
[26] R. Wegrzyk, "Fixed point theorems for multifunctions and their applications to functional equations," *Dissertationes Mathematicae*, vol. 201, pp. 1–28, 1982.
[27] H. Covitz and S. B. Nadler Jr., "Multi-valued contraction mappings in generalized metric spaces," *Israel Journal of Mathematics*, vol. 8, no. 1, pp. 5–11, 1970.
[28] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics 580, Springer, Heidelberg, Germany, 1977.
[29] B. Ahmad and S. K. Ntouyas, "Some existence results for boundary value problems of fractional differential inclusions with non-separated boundary conditions," *Electronic Journal of Qualitative Theory of Differential Equations*, vol. 71, pp. 1–17, 2010.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 631412, 19 pages
doi:10.1155/2011/631412

Research Article

Boundary-Value Problems for Weakly Nonlinear Delay Differential Systems

A. Boichuk,[1,2] J. Diblík,[3,4] D. Khusainov,[5] and M. Růžičková[2]

[1] *Institute of Mathematics, National Academy of Sciences of Ukraine, Tereshchenkovskaya Street 3, 01601 Kyiv, Ukraine*
[2] *Department of Mathematics, University of Žilina, Univerzitná 8215/1, 01026 Žilina, Slovakia*
[3] *Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Brno University of Technology, Veveří 331/95, 602 00 Brno, Czech Republic*
[4] *Department of Mathematics, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technická 10, 616 00 Brno, Czech Republic*
[5] *Department of Complex System Modeling, Faculty of Cybernetics, Taras, Shevchenko National University of Kyiv, Vladimirskaya Street 64, 01033 Kyiv, Ukraine*

Correspondence should be addressed to A. Boichuk, boichuk@imath.kiev.ua

Received 30 January 2011; Accepted 31 March 2011

Academic Editor: Elena Braverman

Conditions are derived of the existence of solutions of nonlinear boundary-value problems for systems of n ordinary differential equations with constant coefficients and single delay (in the linear part) and with a finite number of measurable delays of argument in nonlinearity: $\dot{z}(t) = Az(t-\tau) + g(t) + \varepsilon Z(z(h_i(t), t, \varepsilon)$, $t \in [a,b]$, assuming that these solutions satisfy the initial and boundary conditions $z(s) := \psi(s)$ if $s \notin [a,b]$, $\ell z(\cdot) = \alpha \in \mathbb{R}^m$. The use of a delayed matrix exponential and a method of pseudoinverse by Moore-Penrose matrices led to an *explicit* and *analytical* form of sufficient conditions for the existence of solutions in a given space and, moreover, to the construction of an iterative process for finding the solutions of such problems in a general case when the number of boundary conditions (defined by a linear vector functional ℓ) does not coincide with the number of unknowns in the differential system with a single delay.

1. Introduction

First, we derive some auxiliary results concerning the theory of differential equations with delay. Consider a system of linear differential equations with concentrated delay

$$\dot{z}(t) - A(t)z(h_0(t)) = g(t) \quad \text{if } t \in [a,b], \tag{1.1}$$

to Contents, page 1

assuming that

$$z(s) := \psi(s) \quad \text{if } s \notin [a, b], \tag{1.2}$$

where A is an $n \times n$ real matrix and g is an n-dimensional real column-vector with components in the space $L_p[a, b]$ (where $p \in [1, \infty)$) of functions summable on $[a, b]$; the delay $h_0(t) \leq t$ is a function $h_0 : [a, b] \to \mathbb{R}$ measurable on $[a, b]$; $\psi : \mathbb{R} \setminus [a, b] \to \mathbb{R}^n$ is a given function. Using the denotations

$$(S_{h_0} z)(t) := \begin{cases} z(h_0(t)) & \text{if } h_0(t) \in [a, b], \\ \theta & \text{if } h_0(t) \notin [a, b], \end{cases} \tag{1.3}$$

$$\psi^{h_0}(t) := \begin{cases} \theta & \text{if } h_0(t) \in [a, b], \\ \psi(h_0(t)) & \text{if } h_0(t) \notin [a, b], \end{cases} \tag{1.4}$$

where θ is an n-dimensional zero column-vector and assuming $t \in [a, b]$, it is possible to rewrite (1.1), (1.2) as

$$(Lz)(t) := \dot{z}(t) - A(t)(S_{h_0} z)(t) = \varphi(t), \quad t \in [a, b], \tag{1.5}$$

where φ is an n-dimensional column-vector defined by the formula

$$\varphi(t) := g(t) + A(t)\psi^{h_0}(t) \in L_p[a, b]. \tag{1.6}$$

We will investigate (1.5) assuming that the operator L maps a Banach space $D_p[a, b]$ of absolutely continuous functions $z : [a, b] \to \mathbb{R}^n$ into a Banach space $L_p[a, b]$ $(1 \leq p < \infty)$ of functions $\varphi : [a, b] \to \mathbb{R}^n$ summable on $[a, b]$; the operator S_{h_0} maps the space $D_p[a, b]$ into the space $L_p[a, b]$. Transformations (1.3), (1.4) make it possible to add the initial function $\psi(s)$, $s < a$ to nonhomogeneity generating an additive and homogeneous operation not depending on ψ and without the classical assumption regarding the continuous connection of solution $z(t)$ with the initial function $\psi(t)$ at the point $t = a$.

A solution of differential system (1.5) is defined as an n-dimensional column vector-function $z \in D_p[a, b]$, absolutely continuous on $[a, b]$, with a derivative $\dot{z} \in L_p[a, b]$ satisfying (1.5) almost everywhere on $[a, b]$.

Such approach makes it possible to apply well-developed methods of linear functional analysis to (1.5) with a linear and bounded operator L. It is well-known (see: [1, 2]) that a nonhomogeneous operator equation (1.5) with delayed argument is solvable in the space $D_p[a, b]$ for an arbitrary right-hand side $\varphi \in L_p[a, b]$ and has an n-dimensional family of solutions ($\dim \ker L = n$) in the form

$$z(t) = X(t)c + \int_a^b K(t, s)\varphi(s)\,ds \quad \forall c \in \mathbb{R}^n, \tag{1.7}$$

to Contents, page 1

where the kernel $K(t,s)$ is an $n \times n$ Cauchy matrix defined in the square $[a,b] \times [a,b]$ being, for every fixed $s \leq t$, a solution of the matrix Cauchy problem

$$(LK(\cdot,s))(t) := \frac{\partial K(t,s)}{\partial t} - A(t)(S_{h_0}K(\cdot,s))(t) = \Theta, \quad K(s,s) = I, \tag{1.8}$$

where $K(t,s) \equiv \Theta$ if $a \leq t < s \leq b$, Θ is $n \times n$ null matrix and I is $n \times n$ identity matrix. A fundamental $n \times n$ matrix $X(t)$ for the homogeneous ($\varphi \equiv \theta$) equation (1.5) has the form $X(t) = K(t,a)$, $X(a) = I$ [2]. Throughout the paper, we denote by Θ_s an $s \times s$ null matrix if $s \neq n$, by $\Theta_{s,p}$ an $s \times p$ null matrix, by I_s an $s \times s$ identity matrix if $s \neq n$, and by θ_s an s-dimensional zero column-vector if $s \neq n$.

A serious disadvantage of this approach, when investigating the above-formulated problem, is the necessity to find the Cauchy matrix $K(t,s)$ [3, 4]. It exists but, as a rule, can only be found numerically. Therefore, it is important to find systems of differential equations with delay such that this problem can be solved directly. Below we consider the case of a system with so-called single delay [5]. In this case, the problem of how to construct the Cauchy matrix is successfully solved *analytically* due to a delayed matrix exponential defined below.

1.1. A Delayed Matrix Exponential

Consider a Cauchy problem for a linear nonhomogeneous differential system with constant coefficients and with a single delay τ

$$\dot{z}(t) = Az(t-\tau) + g(t), \tag{1.9}$$

$$z(s) = \psi(s), \quad \text{if } s \in [-\tau, 0], \tag{1.10}$$

with an $n \times n$ constant matrix A, $g : [0,\infty) \to \mathbb{R}^n$, $\psi : [-\tau, 0] \to \mathbb{R}^n$, $\tau > 0$ and an unknown vector-solution $z : [-\tau,\infty) \to \mathbb{R}^n$. Together with a nonhomogeneous problem (1.9), (1.10), we consider a related homogeneous problem

$$\begin{aligned} \dot{z}(t) &= Az(t-\tau), \\ z(s) &= \psi(s), \quad \text{if } s \in [-\tau, 0]. \end{aligned} \tag{1.11}$$

to Contents, page 1

Denote by e_τ^{At} a matrix function called a delayed matrix exponential (see [5]) and defined as

$$e_\tau^{At} := \begin{cases} \Theta & \text{if } -\infty < t < -\tau, \\ I & \text{if } -\tau \le t < 0, \\ I + A\dfrac{t}{1!} & \text{if } 0 \le t < \tau, \\ I + A\dfrac{t}{1!} + A^2\dfrac{(t-\tau)^2}{2!} & \text{if } \tau \le t < 2\tau, \\ \cdots & \\ I + A\dfrac{t}{1!} + \cdots + A^k\dfrac{(t-(k-1)\tau)^k}{k!} & \text{if } (k-1)\tau \le t < k\tau, \\ \cdots . & \end{cases} \tag{1.12}$$

This definition can be reduced to the following expression:

$$e_\tau^{At} = \sum_{n=0}^{[t/\tau]+1} A^n \frac{(t-(n-1)\tau)^n}{n!}, \tag{1.13}$$

where $[t/\tau]$ is the greatest integer function. The delayed matrix exponential equals the unit matrix I on $[-\tau, 0]$ and represents a fundamental matrix of a homogeneous system with single delay. Thus, the delayed matrix exponential solves the Cauchy problem for a homogeneous system (1.11), satisfying the unit initial conditions

$$z(s) = \varphi(s) \equiv e_\tau^{As} = I \quad \text{if } -\tau \le s \le 0, \tag{1.14}$$

and the following statement holds (see, e.g., [5], [6, Remark 1], [7, Theorem 2.1]).

Lemma 1.1. *A solution of a Cauchy problem for a nonhomogeneous system with single delay* (1.9), *satisfying a constant initial condition*

$$z(s) = \varphi(s) = c \in \mathbb{R}^n \quad \textit{if } s \in [-\tau, 0] \tag{1.15}$$

has the form

$$z(t) = e_\tau^{A(t-\tau)}c + \int_0^t e_\tau^{A(t-\tau-s)} g(s)ds. \tag{1.16}$$

The delayed matrix exponential was applied, for example, in [6, 7] to investigation of boundary value problems of diffferential systems with a single delay and in [8] to investigation of the stability of linear perturbed systems with a single delay.

to Contents, page 1

1.2. Fredholm Boundary-Value Problem

Without loss of generality, let $a = 0$ and, with a view of the above, the problem (1.9), (1.10) can be transformed ($h_0(t) := t - \tau$) to an equation of the type (1.1) (see (1.5))

$$\dot{z}(t) - A(S_{h_0}z)(t) = \varphi(t), \quad t \in [0, b], \tag{1.17}$$

where, in accordance with (1.3),(1.4)

$$\begin{gathered}(S_{h_0}z)(t) = \begin{cases} z(t-\tau) & \text{if } t-\tau \in [0,b], \\ \theta & \text{if } t-\tau \notin [0,b], \end{cases} \\ \varphi(t) = \ g(t) + A\ \psi^{h_0}(t)\ \in\ L_p[0,b], \\ \psi^{h_0}(t) = \begin{cases} \theta & \text{if } t-\tau \in [0,b], \\ \psi(t-\tau) & \text{if } t-\tau \notin [0,b]. \end{cases}\end{gathered} \tag{1.18}$$

A general solution of problem (1.17) for a nonhomogeneous system with single delay and zero initial data has the form (1.7)

$$z(t) = X(t)c + \int_0^b K(t,s)\varphi(s)ds \quad \forall c \in \mathbb{R}^n, \tag{1.19}$$

where, as can easily be verified (in view of the above-defined delayed matrix exponential) by substituting into (1.17),

$$X(t) = e_\tau^{A(t-\tau)}, \qquad X(0) = e_\tau^{-A\tau} = I \tag{1.20}$$

is a normal fundamental matrix of the homogeneous system related to (1.17) (or (1.9)) with initial data $X(0) = I$, and the Cauchy matrix $K(t,s)$ has the form

$$\begin{gathered}K(t,s) = e_\tau^{A(t-\tau-s)} \quad \text{if } 0 \le s < t \le b, \\ K(t,s) \equiv \Theta \quad \text{if } 0 \le t < s \le b.\end{gathered} \tag{1.21}$$

Obviously

$$K(t,0) = e_\tau^{A(t-\tau)} = X(t), \qquad K(0,0) = e_\tau^{A(-\tau)} = X(0) = I, \tag{1.22}$$

and, therefore, the initial problem (1.17) for systems of ordinary differential equations with constant coefficients and single delay has an n-parametric family of linearly independent solutions (1.16).

Now, we will deal with a general boundary-value problem for system (1.17). Using the results [2, 9], it is easy to derive statements for a general boundary-value problem if the

to Contents, page 1

number m of boundary conditions does not coincide with the number n of unknowns in a differential system with single delay.

We consider a boundary-value problem

$$\begin{aligned} \dot{z}(t) - Az(t-\tau) &= g(t), \quad t \in [0,b], \\ z(s) &:= \psi(s), \quad s \notin [0,b], \end{aligned} \tag{1.23}$$

assuming that

$$\ell z(\cdot) = \alpha \in \mathbb{R}^m, \tag{1.24}$$

or, using (1.18), its equivalent form

$$\begin{aligned} \dot{z}(t) - A(S_{h_0}z)(t) &= \varphi(t), \quad t \in [0,b], \\ \ell z(\cdot) &= \alpha \in \mathbb{R}^m, \end{aligned} \tag{1.25}$$

where α is an m-dimensional constant vector-column ℓ is an m-dimensional linear vector-functional defined on the space $D_p[0,b]$ of an n-dimensional vector-functions

$$\ell = \text{col}\,(\ell_1, \ldots, \ell_m) : D_p[0,b] \longrightarrow \mathbb{R}^m, \qquad \ell_i : D_p[0,b] \longrightarrow \mathbb{R}, \quad i = 1, \ldots, m, \tag{1.26}$$

absolutely continuous on $[0,b]$. Such problems for functional-differential equations are of Fredholm's type (see, e.g., [1, 2]). In order to formulate the following result, we need several auxiliary abbreviations. We set

$$Q := \ell X(\cdot) = \ell e_\tau^{A(\cdot-\tau)}. \tag{1.27}$$

We define an $n \times n$-dimensional matrix (orthogonal projection)

$$P_Q := I - Q^+Q, \tag{1.28}$$

projecting space $\mathbb{R}^n$ to ker Q of the matrix Q.

Moreover, we define an $m \times m$-dimensional matrix (orthogonal projection)

$$P_{Q^*} := I_m - QQ^+, \tag{1.29}$$

projecting space $\mathbb{R}^m$ to ker Q^* of the transposed matrix $Q^* = Q^T$, where I_m is an $m \times m$ identity matrix and Q^+ is an $n \times m$-dimensional matrix pseudoinverse to the $m \times n$-dimensional matrix Q. Denote $d := \text{rank}\, P_{Q^*}$ and $n_1 := \text{rank}\, Q = \text{rank}\, Q^*$. Since

$$\text{rank}\, P_{Q^*} = m - \text{rank}\, Q^*, \tag{1.30}$$

we have $d = m - n_1$.

to Contents, page 1

We will denote by $P_{Q_d^*}$ an $d \times m$-dimensional matrix constructed from d linearly independent rows of the matrix P_{Q^*}. Denote $r := \operatorname{rank} P_Q$. Since

$$\operatorname{rank} P_Q = n - \operatorname{rank} Q, \tag{1.31}$$

we have $r = n - n_1$. By P_{Q_r} we will denote an $n \times r$-dimensional matrix constructed from r linearly independent columns of the matrix P_Q. Finally, we define

$$X_r(t) := X(t)P_{Q_r}, \tag{1.32}$$

and a generalized Green operator

$$(G\varphi)(t) := \int_0^b G(t,s)\varphi(s)ds, \tag{1.33}$$

where

$$G(t,s) := K(t,s) - e_\tau^{A(t-\tau)} Q^+ \ell K(\cdot, s) \tag{1.34}$$

is a generalized Green matrix corresponding to the boundary-value problem (1.25) (the Cauchy matrix $K(t,s)$ has the form (1.21)).

In [6, Theorem 4], the following result (formulating the necessary and sufficient conditions of solvability and giving representations of the solutions $z \in D_p[0,b]$, $\dot{z} \in L_p[0,b]$ of the boundary-value problem (1.25) in an *explicit analytical* form) is proved.

Theorem 1.2. *If $n_1 \le \min(m,n)$, then:*

(i) *the homogeneous problem*

$$\begin{aligned} \dot{z}(t) - A(S_{h_0}z)(t) &= \theta, \quad t \in [0,b], \\ \ell z(\cdot) &= \theta_m \in \mathbb{R}^m \end{aligned} \tag{1.35}$$

corresponding to problem (1.25) *has exactly r linearly independent solutions*

$$z(t, c_r) = X_r(t)c_r = e_\tau^{A(t-\tau)} P_{Q_r} c_r \in D_p[0,b], \tag{1.36}$$

(ii) *nonhomogeneous problem* (1.25) *is solvable in the space $D_p[0,b]$ if and only if $\varphi \in L_p[0,b]$ and $\alpha \in \mathbb{R}^m$ satisfy d linearly independent conditions*

$$P_{Q_d^*} \cdot \left(\alpha - \ell \int_0^b K(\cdot, s)\varphi(s)ds \right) = \theta_d, \tag{1.37}$$

to Contents, page 1

(iii) *in that case the nonhomogeneous problem* (1.25) *has an* r*-dimensional family of linearly independent solutions represented in an analytical form*

$$z(t) = z_0(t, c_r) := X_r(t)c_r + \big(G\varphi\big)(t) + X(t)Q^{+}\alpha \quad \forall c_r \in \mathbb{R}^r. \tag{1.38}$$

2. Perturbed Weakly Nonlinear Boundary Value Problems

As an example of applying Theorem 1.2, we consider a problem of the branching of solutions $z : [0, b] \to \mathbb{R}^n$, $b > 0$ of systems of nonlinear ordinary differential equations with a small parameter ε and with a finite number of measurable delays $h_i(t)$, $i = 1, 2, \ldots, k$ of argument of the form

$$\dot{z}(t) = Az(t-\tau) + g(t) + \varepsilon Z(z(h_i(t)), t, \varepsilon), \quad t \in [0, b],\ h_i(t) \le t, \tag{2.1}$$

satisfying the initial and boundary conditions

$$z(s) = \psi(s), \quad \text{if } s < 0, \qquad \ell z(\cdot) = \alpha, \quad \alpha \in \mathbb{R}^m, \tag{2.2}$$

and such that its solution $z = z(t, \varepsilon)$, satisfying

$$\begin{aligned} &z(\cdot, \varepsilon) \in D_p[0, b],\\ &\dot{z}(\cdot, \varepsilon) \in L_p[0, b],\\ &z(t, \cdot) \in C[0, \varepsilon_0], \end{aligned} \tag{2.3}$$

for a sufficiently small $\varepsilon_0 > 0$, for $\varepsilon = 0$, turns into one of the generating solutions (1.38); that is, $z(t, 0) = z_0(t, c_r)$ for a $c_r \in \mathbb{R}^r$. We assume that the $n \times 1$ vector-operator Z satisfies

$$\begin{aligned} &Z(\cdot, t, \varepsilon) \in C^1\left[\|z - z_0\| \le q\right],\\ &Z(z(h_i(t)), \cdot, \varepsilon) \in L_p[0, b],\\ &Z(z(h_i(t)), t, \cdot) \in C[0, \varepsilon_0], \end{aligned} \tag{2.4}$$

where $q > 0$ is sufficiently small. Using denotations (1.3), (1.4), and (1.6), it is easy to show that the perturbed nonlinear boundary value problem (2.1), (2.2) can be rewritten in the form

$$\dot{z}(t) = A(S_{h_0}z)(t) + \varepsilon Z((S_h z)(t), t, \varepsilon) + \varphi(t), \quad \ell z(\cdot) = \alpha,\ t \in [0, b]. \tag{2.5}$$

In (2.5), A is an $n \times n$ constant matrix, $h_0 : [0, b] \to \mathbb{R}$ is a single delay defined by $h_0(t) := t - \tau$, $\tau > 0$,

$$(S_h z)(t) = \operatorname{col}[(S_{h_1}z)(t), \ldots, (S_{h_k}z)(t)] \tag{2.6}$$

to Contents, page 1

is an N-dimensional column vector, where $N = nk$, and φ is an n-dimensional column vector given by

$$\varphi(t) = g(t) + A\ \varphi^{h_0}(t). \tag{2.7}$$

The operator S_h maps the space D_p into the space

$$L_p^N = \underbrace{L_p \times \cdots \times L_p}_{k\text{-times}}, \tag{2.8}$$

that is, $S_h : D_p \to L_p^N$. Using denotation (1.3) for the operator $S_{h_i} : D_p \to L_p$, $i = 1, \dots, k$, we have the following representation:

$$(S_{h_i} z)(t) = \int_0^b \chi_{h_i}(t, s)\dot{z}(s)ds + \chi_{h_i}(t, 0)z(0), \tag{2.9}$$

where

$$\chi_{h_i}(t, s) = \begin{cases} 1, & \text{if } (t, s) \in \Omega_i, \\ 0, & \text{if } (t, s) \notin \Omega_i \end{cases} \tag{2.10}$$

is the characteristic function of the set

$$\Omega_i := \{(t, s) \in [0, b] \times [0, b] : 0 \le s \le h_i(t) \le b\}. \tag{2.11}$$

Assume that the generating boundary value problem

$$\dot{z}(t) = A(S_{h_0} z)(t) + \varphi(t), \quad lz = \alpha, \tag{2.12}$$

being a particular case of (2.5) for $\varepsilon = 0$, has solutions for nonhomogeneities $\varphi \in L_p[0, b]$ and $\alpha \in \mathbb{R}^m$ that satisfy conditions (1.37). In such a case, by Theorem 1.2, the problem (2.12) possesses an r-dimensional family of solutions of the form (1.38).

Problem 1. Below, we consider the following problem: derive the necessary and sufficient conditions indicating when solutions of (2.5) turn into solutions (1.38) of the boundary value problem (2.12) for $\varepsilon = 0$.

Using the theory of generalized inverse operators [2], it is possible to find conditions for the solutions of the boundary value problem (2.5) to be branching from the solutions of (2.5) with $\varepsilon = 0$. Below, we formulate statements, solving the above problem. As compared with an earlier result [10, page 150], the present result is derived in an *explicit analytical* form. The progress was possible by using the delayed matrix exponential since, in such a case, all the necessary calculations can be performed to the full.

to Contents, page 1

Theorem 2.1 (necessary condition). *Consider the system* (2.1)*; that is,*

$$\dot{z}(t) = Az(t-\tau) + g(t) + \varepsilon Z(z(h_i(t)), t, \varepsilon), \quad t \in [0, b], \tag{2.13}$$

where $h_i(t) \le t$, $i = 1, \ldots, k$, with the initial and boundary conditions (2.2)*; that is,*

$$z(s) = \psi(s), \quad \text{if } s < 0 < b, \qquad \ell z(\cdot) = \alpha \ \in \mathbb{R}^m, \tag{2.14}$$

and assume that, for nonhomogeneities

$$\varphi(t) = \ g(t) + A\ \psi^{h_0}(t) \ \in L_p[0, b], \tag{2.15}$$

and for $\alpha \in \mathbb{R}^m$, the generating boundary value problem

$$\dot{z}(t) = A(S_{h_0}z)(t) + \varphi(t), \quad \ell z(\cdot) = \alpha, \tag{2.16}$$

corresponding to the problem (1.25)*, has exactly an r-dimensional family of linearly independent solutions of the form* (1.38)*. Moreover, assume that the boundary value problem* (2.13)*,* (2.14) *has a solution $z(t, \varepsilon)$ which, for $\varepsilon = 0$, turns into one of solutions $z_0(t, c_r)$ in* (1.38) *with a vector-constant $c_r := c_r^0 \in \mathbb{R}^r$.*

Then, the vector c_r^0 satisfies the equation

$$F\left(c_r^0\right) := \int_0^b H(s)Z\Big((S_h z_0)\left(s, c_r^0\right), s, 0\Big)ds = \theta_d, \tag{2.17}$$

where

$$H(s) := P_{Q_d^*}\ell K(\cdot, s) = P_{Q_d^*}\ell e_\tau^{A(\cdot-\tau-s)}. \tag{2.18}$$

Proof. We consider the nonlinearity in system (2.13), that is, the term $\varepsilon Z(z(h_i(t)), t, \varepsilon)$ as an inhomogeneity, and use Theorem 1.2 assuming that condition (1.37) is satisfied. This gives

$$\int_0^b H(s)Z((S_h z)(s, \varepsilon), s, \varepsilon)ds = \theta_d. \tag{2.19}$$

In this integral, letting $\varepsilon \to 0$, we arrive at the required condition (2.17). □

Corollary 2.2. *For periodic boundary-value problems, the vector-constant $c_r \in \mathbb{R}^r$ has a physical meaning-it is the amplitude of the oscillations generated. For this reason,* (2.17) *is called an equation generating the amplitude [11]. By analogy with the investigation of periodic problems, it is natural to say* (2.17) *is an equation for generating the constants of the boundary value problem* (2.13)*,* (2.14)*.*

to Contents, page 1

If (2.17) *is solvable, then the vector constant* $c_r^0 \in \mathbb{R}^r$ *specifies the generating solution* $z_0(t, c_r^0)$ *corresponding to the solution* $z = z(t, \varepsilon)$ *of the original problem such that*

$$\begin{gathered} z(\cdot, \varepsilon) : [0, b] \longrightarrow \mathbb{R}^n, \\ z(\cdot, \varepsilon) \in D_p[0, b], \\ \dot{z}(\cdot, \varepsilon) \in L_p[0, b], \\ z(t, \cdot) \in C[0, \varepsilon_0], \\ z(t, 0) = z_0\left(t, c_r^0\right). \end{gathered} \tag{2.20}$$

Also, if (2.17) *is unsolvable, the problem* (2.13), (2.14) *has no solution in the analyzed space. Note that, here and in what follows, all expressions are obtained in the real form and hence, we are interested in real solutions of* (2.17), *which can be algebraic or transcendental.*

Sufficient conditions for the existence of solutions of the boundary-value problem (2.13), (2.14) can be derived using results in [10, page 155] and [2]. By changing the variables in system (2.13), (2.14)

$$z(t, \varepsilon) = z_0\left(t, c_r^0\right) + y(t, \varepsilon), \tag{2.21}$$

we arrive at a problem of finding sufficient conditions for the existence of solutions of the problem

$$\dot{y}(t) = A\left(S_{h_0} y\right)(t) + \varepsilon Z\left(S_h\left(z_0 + y\right)(t), t, \varepsilon\right), \quad \ell y = \theta_m, \ t \in [0, b], \tag{2.22}$$

and such that

$$\begin{gathered} y(\cdot, \varepsilon) : [0, b] \longrightarrow \mathbb{R}^n, \\ y(\cdot, \varepsilon) \in D_p[0, b], \\ \dot{y}(\cdot, \varepsilon) \in L_p[0, b], \\ y(t, \cdot) \in C[0, \varepsilon_0], \\ y(t, 0) = \theta. \end{gathered} \tag{2.23}$$

Since the vector function $Z((S_h z)(t), t, \varepsilon)$ is continuously differentiable with respect to z and continuous in ε in the neighborhood of the point

$$(z, \varepsilon) = \left(z_0\left(t, c_r^0\right), 0\right), \tag{2.24}$$

to Contents, page 1

we can separate its linear term as a function depending on y and terms of order zero with respect to ε

$$Z\left(S_h\left(z_0\left(t,c_r^0\right)+y\right),t,\varepsilon\right)=f_0\left(t,c_r^0\right)+A_1(t)(S_hy)(t)+R((S_hy)(t),t,\varepsilon), \tag{2.25}$$

where

$$\begin{gathered} f_0\left(t,c_r^0\right):=Z\left((S_hz_0)\left(t,c_r^0\right),t,0\right), \quad f_0\left(\cdot,c_r^0\right)\in L_p[0,b], \\ A_1(t)=A_1\left(t,c_r^0\right)=\left.\frac{\partial Z(S_hx,t,0)}{\partial S_hx}\right|_{x=z_0(t,c_r^0)}, \quad A_1(\cdot)\in L_p[0,b], \\ R(\theta,t,0)=\theta, \quad \frac{\partial R(\theta,t,0)}{\partial y}=\Theta, \quad R(y,\cdot,\varepsilon)\in L_p[0,b]. \end{gathered} \tag{2.26}$$

We now consider the vector function $Z((S_h(z_0+y))(t),t,\varepsilon)$ in (2.22) as an inhomogeneity and we apply Theorem 1.2 to this system. As the result, we obtain the following representation for the solution of (2.22):

$$y(t,\varepsilon)=X_r(t)c+y^{(1)}(t,\varepsilon). \tag{2.27}$$

In this expression, the unknown vector of constants $c=c(\varepsilon)\in C[0,\varepsilon_0]$ is determined from a condition similar to condition (1.37) for the existence of solution of problem (2.22):

$$B_0c=\int_0^b H(s)\left[A_1(s)\left(S_hy^{(1)}\right)(s,\varepsilon)+R((S_hy)(s,\varepsilon),s,\varepsilon)\right]ds, \tag{2.28}$$

where

$$B_0=\int_0^b H(s)A_1(s)(S_hX_r)(s)ds \tag{2.29}$$

is a $d\times r$ matrix, and

$$H(s):=P_{Q_d^*}\ell K(\cdot,s)=P_{Q_d^*}\ell e_\tau^{A(\cdot-\tau-s)}. \tag{2.30}$$

The unknown vector function $y^{(1)}(t,\varepsilon)$ is determined by using the generalized Green operator as follows:

$$y^{(1)}(t,\varepsilon)=\varepsilon\left(G\left[Z\left(S_h\left(z_0\left(s,c_r^0\right)+y\right),s,\varepsilon\right)\right]\right)(t). \tag{2.31}$$

to Contents, page 1

Let $P_{N(B_0)}$ be an $r \times r$ matrix orthoprojector $\mathbb{R}^r \to N(B_0)$, and let $P_{N(B_0^*)}$ be a $d \times d$ matrix-orthoprojector $\mathbb{R}^d \to N(B_0^*)$. Equation (2.28) is solvable with respect to $c \in \mathbb{R}^r$ if and only if

$$P_{N(B_0^*)} \int_0^b H(s)\Big[A_1(s)\Big(S_h y^{(1)}\Big)(s,\varepsilon) + R\big((S_h y)(s,\varepsilon), s, \varepsilon\big)\Big] ds = \theta_d. \tag{2.32}$$

For

$$P_{N(B_0^*)} = \Theta_d, \tag{2.33}$$

the last condition is always satisfied and (2.28) is solvable with respect to $c \in \mathbb{R}^r$ up to an arbitrary vector constant $P_{N(B_0)}c \in \mathbb{R}^r$ from the null space of the matrix B_0

$$c = B_0^+ \int_0^b H(s)\Big[A_1(s)\Big(S_h y^{(1)}\Big)(s,\varepsilon) + R\big((S_h y)(s,\varepsilon), s, \varepsilon\big)\Big] ds + P_{N(B_0)}c. \tag{2.34}$$

To find a solution $y = y(t,\varepsilon)$ of (2.28) such that

$$\begin{gathered} y(\cdot,\varepsilon) : [0,b] \longrightarrow R^n, \\ y(\cdot,\varepsilon) \in D_p[0,b], \\ \dot{y}(\cdot,\varepsilon) \in L_p[0,b], \\ y(t,\cdot) \in C[0,\varepsilon_0], \\ y(t,0) = \theta, \end{gathered} \tag{2.35}$$

it is necessary to solve the following operator system:

$$\begin{gathered} y(t,\varepsilon) = X_r(t)c + y^{(1)}(t,\varepsilon), \\ c = B_0^+ \int_0^b H(s)\Big[A_1(s)\Big(S_h y^{(1)}\Big)(s,\varepsilon) + R\big((S_h y)(s,\varepsilon), s, \varepsilon\big)\Big] ds, \\ y^{(1)}(t,\varepsilon) = \varepsilon G\Big[Z\Big(S_h\Big(z_0\Big(s, c_r^0\Big) + y\Big), s, \varepsilon\Big)\Big](t). \end{gathered} \tag{2.36}$$

The operator system (2.36) belongs to the class of systems solvable by the method of simple iterations, convergent for sufficiently small $\varepsilon \in [0,\varepsilon_0]$ (see [10, page 188]). Indeed, system (2.36) can be rewritten in the form

$$u = L^{(1)}u + Fu, \tag{2.37}$$

to Contents, page 1

where $u = \operatorname{col}\,(y(t,\varepsilon), c(\varepsilon), y^{(1)}(t,\varepsilon))$ is a $(2n+r)$-dimensional column vector, $L^{(1)}$ is a linear operator

$$L^{(1)} := \begin{pmatrix} \Theta & X_r(t) & I \\ \Theta_{r,n} & \Theta_{r,r} & L_1 \\ \Theta & \Theta_{n,r} & \Theta \end{pmatrix}, \tag{2.38}$$

where

$$L_1(*) = B_0^+ \int_0^b H(s) A_1(s)(*) ds, \tag{2.39}$$

and F is a nonlinear operator

$$Fu := \begin{pmatrix} \theta \\ B_0^+ \int_0^b H(s) R\big((S_h y)(s,\varepsilon), s, \varepsilon\big) ds \\ \varepsilon\big(G\big[Z\big((S_h z_0)(s, c_r^0), s, 0\big) + A_1(s)\big(S_h y\big)(s,\varepsilon) + R\big((S_h y)(s,\varepsilon), s, \varepsilon\big)\big]\big)(t) \end{pmatrix}. \tag{2.40}$$

In view of the structure of the operator $L^{(1)}$ containing zero blocks on and below the main diagonal, the inverse operator

$$\left(I_{2n+r} - L^{(1)}\right)^{-1} \tag{2.41}$$

exists. System (2.37) can be transformed into

$$u = Su, \tag{2.42}$$

where

$$S := \left(I_{2n+r} - L^{(1)}\right)^{-1} F \tag{2.43}$$

is a contraction operator in a sufficiently small neighborhood of the point

$$(z, \varepsilon) = \left(z_0\left(t, c_r^0\right), 0\right). \tag{2.44}$$

Thus, the solvability of the last operator system can be established by using one of the existing versions of the fixed-point principles [12] applicable to the system for sufficiently small $\varepsilon \in [0, \varepsilon_0]$. It is easy to prove that the sufficient condition $P_{N(B_0^*)} = \Theta_d$ for the existence of solutions of the boundary value problem (2.13), (2.14) means that the constant $c_r^0 \in \mathbb{R}^r$ of the equation for generating constant (2.17) is a simple root of equation (2.17) [2]. By using the method of simple iterations, we can find the solution of the operator system and hence the solution of the original boundary value problem (2.13), (2.14). Now, we arrive at the following theorem.

to Contents, page 1

Theorem 2.3 (sufficient condition). *Assume that the boundary value problem* (2.13), (2.14) *satisfies the conditions listed above and the corresponding linear boundary value problem* (1.25) *has an r-dimensional family of linearly independent solutions of the form* (1.38). *Then, for any simple root $c_r = c_r^0 \in \mathbb{R}^r$ of the equation for generating the constants* (2.17), *there exist at least one solution of the boundary value problem* (2.13), (2.14). *The indicated solution $z(t,\varepsilon)$ is such that*

$$\begin{aligned} z(\cdot,\varepsilon) &\in D_p[0,b], \\ \dot{z}(\cdot,\varepsilon) &\in L_p[0,b], \\ z(t,\cdot) &\in C[0,\varepsilon_0], \end{aligned} \tag{2.45}$$

and, for $\varepsilon = 0$, turns into one of the generating solutions (1.38) *with a constant $c_r^0 \in \mathbb{R}^r$; that is, $z(t,0) = z_0(t,c_r^0)$. This solution can be found by the method of simple iterations, which is convergent for a sufficiently small $\varepsilon \in [0,\varepsilon_0]$.*

Corollary 2.4. *If the number n of unknown variables is equal to the number m of boundary conditions (and hence $r = d$), the boundary value problem* (2.13), (2.14) *has a unique solution. In such a case, the problems considered for functional-differential equations are of Fredholm's type with a zero index. By using the procedure proposed in [2] with some simplifying assumptions, we can generalize the proposed method to the case of multiple roots of equation* (2.17) *to determine sufficient conditions for the existence of solutions of the boundary-value problem* (2.13), (2.14).

3. Example

We will illustrate the above proved theorems on the example of a weakly perturbed linear boundary value problem. Consider the following simplest boundary value problem-a periodic problem for the delayed differential equation:

$$\begin{gathered} \dot{z}(t) = z(t-\tau) + \varepsilon \sum_{i=1}^{k} B_i(t) z(h_i(t)) + g(t), \quad t \in (0,T], \\ z(s) = \psi(s), \quad \text{if } s < 0, \\ z(0) = z(T), \end{gathered} \tag{3.1}$$

where $0 < \tau, T = \text{const}$, B_i are $n \times n$ matrices, $B_i, g \in L_p[0,T]$, $\psi : \mathbb{R}^1 \setminus (0,T] \to \mathbb{R}^n$, $h_i(t) \le t$ are measurable functions. Using the symbols S_{h_i} and ψ^{h_i} (see (1.3), (1.4), (2.9)), we arrive at the following operator system:

$$\begin{gathered} \dot{z}(t) = z(t-\tau) + \varepsilon B(t)(S_h z)(t) + \varphi(t,\varepsilon), \\ \ell z := z(0) - z(T) = \theta_n, \end{gathered} \tag{3.2}$$

to Contents, page 1

where $B(t) := (B_1(t), \ldots, B_k(t))$ is an $n \times N$ matrix ($N = nk$), and

$$\varphi(t, \varepsilon) := g(t) + \psi^{h_0}(t) + \varepsilon \sum_{i=1}^{k} B_i(t) \psi^{h_i}(t) \in L_p[0, T]. \tag{3.3}$$

We will consider the simplest case with $T \leq \tau$. Utilizing the delayed matrix exponential, it can be easily verified that in this case, the matrix

$$X(t) = e_\tau^{I(t-\tau)} = I \tag{3.4}$$

is a normal fundamental matrix for the homogeneous generating system

$$\dot{z}(t) = z(t - \tau). \tag{3.5}$$

Then,

$$\begin{gathered}
Q := \ell X(\cdot) = e_\tau^{-I\tau} - e_\tau^{I(T-\tau)} = \theta_n, \\
P_Q = P_{Q^*} = I, \quad (r = n, d = m = n), \\
K(t, s) = \begin{cases} e_\tau^{I(t-\tau-s)} = I, & 0 \leq s \leq t \leq T, \\ \Theta, & s > t, \end{cases} \\
\ell K(\cdot, s) = K(0, s) - K(T, s) = -I, \\
H(\tau) = P_{Q^*} \ell K(\cdot, s) = -I, \\
(S_{h_i} I)(t) = \chi_{h_i}(t, 0) \cdot I = I \cdot \begin{cases} 1, & \text{if } 0 \leq h_i(t) \leq T, \\ 0, & \text{if } h_i(t) < 0. \end{cases}
\end{gathered} \tag{3.6}$$

To illustrate the theorems proved above, we will find the conditions for which the boundary value problem (3.1) has a solution $z(t, \varepsilon)$ that, for $\varepsilon = 0$, turns into one of solutions (1.38) $z_0(t, c_r)$ of the generating problem. In contrast to the previous works [7, 9], we consider the case when the unperturbed boundary-value problem

$$\begin{gathered}
\dot{z}(t) = z(t - \tau) + \varphi(t, 0), \\
z(0) = z(T)
\end{gathered} \tag{3.7}$$

has an n-parametric family of linear-independent solutions of the form(1.38)

$$z := z_0(t, c_n) = c_n + (G\varphi)(t), \quad \forall c_n \in \mathbb{R}^n. \tag{3.8}$$

For this purpose, it is necessary and sufficient for the vector function

$$\varphi(t) = g(t) + \psi^{h_0}(t) \tag{3.9}$$

to Contents, page 1

to satisfy the condition of type (1.37)

$$\int_0^T H(s)\varphi(s)\ ds = -\int_0^T \varphi(s)\ ds = \theta_n. \tag{3.10}$$

Then, according to the Theorem 2.1, the constant $c_n = c_n^0 \in \mathbb{R}^n$ must satisfy (2.17), that is, the equation

$$F\left(c_n^0\right) := \int_0^T H(s)Z\left((S_h z_0)\left(s, c_n^0\right), s, 0\right)ds = \theta_n, \tag{3.11}$$

which in our case is a linear algebraic system

$$B_0 c_n^0 = -\int_0^T B(s)(S_h(G\varphi))(s)ds, \tag{3.12}$$

with the $n \times n$ matrix B_0 in the form

$$\begin{aligned} B_0 &= \int_0^T H(s)B(s)(S_h I)(s)ds \\ &= -\int_0^T \sum_{i=1}^k B_i(s)(S_{h_i} I)(s)ds = -\sum_{i=1}^k \int_0^T B_i(s)\chi_{h_i}(s,0)ds. \end{aligned} \tag{3.13}$$

According to Corollary 2.4, if $\det B_0 \neq 0$, the problem (3.1) for the case $T \leq \tau$ has a unique solution $z(t,\varepsilon)$ with the properties

$$\begin{gathered} z(\cdot,\varepsilon) \in D_p^n[0,T], \\ \dot{z}(\cdot,\varepsilon) \in L_p^n[0,T], \\ z(t,\cdot) \in C[0,\varepsilon_0], \\ z(t,0) = z_0\left(t, c_n^0\right), \end{gathered} \tag{3.14}$$

for $g \in L_p[0,T]$, $\varphi(t) \in L_p[0,T]$, and for measurable delays h_i that which satisfy the criterion (3.10) of the existence of a generating solution where

$$c_n^0 = -B_0^+ \int_0^T B(s)(S_h(G\varphi))(s)ds. \tag{3.15}$$

A solution $z(t,\varepsilon)$ of the boundary value problem (3.1) can be found by the convergent method of simple iterations (see Theorem 2.3).

to Contents, page 1

If, for example, $h_i(t) = t - \Delta_i$, where $0 < \Delta_i = \text{const} < T$, $i = 1, \dots, k$, then

$$\chi_{h_i}(t,0) = \begin{cases} 1 & \text{if } 0 \le h_i(t) = t - \Delta_i \le T, \\ 0 & \text{if } h_i(t) = t - \Delta_i < 0, \end{cases} = \begin{cases} 1 & \text{if } \Delta_i \le t \le T + \Delta_{i,} \\ 0, & \text{if } t < \Delta_i. \end{cases} \tag{3.16}$$

The $n \times n$ matrix B_0 can be rewritten in the form

$$\begin{aligned} B_0 &= \int_0^T H(s) \sum_{i=1}^k B_i(s) \chi_{h_i}(s,0) d\tau \\ &= -\sum_{i=1}^k \int_0^T B_i(s) \chi_{h_i}(s,0) ds = -\sum_{i=1}^k \int_{\Delta_i}^T B_i(s) ds, \end{aligned} \tag{3.17}$$

and the unique solvability condition of the boundary value problem (3.1) takes the form

$$\det \left[\sum_{i=1}^k \int_{\Delta_i}^T B_i(s) ds \right] \ne 0. \tag{3.18}$$

It is easy to see that if the vector function $Z(z(h_i(t)), t, \varepsilon)$ is nonlinear in z, for example as a square, then (3.11) generating the constants will be a square-algebraic system and, in this case, the boundary value problem (3.1) can have two solutions branching from the point $\varepsilon = 0$.

Acknowledgments

The first and the fourth authors were supported by the Grant no. 1/0090/09 of the Grant Agency of Slovak Republic (VEGA) and by the project APVV-0700-07 of Slovak Research and Development Agency. The second author was supported by the Grant no. P201/11/0768 of Czech Grant Agency, by the Council of Czech Government MSM 0021630503 and by the Project FEKT/FSI-S-11-1-1159. The third author was supported by the Project no. M/34-2008 of Ukrainian Ministry of Education, Ukraine.

References

[1] N. V. Azbelev and V. P. Maksimov, “Equations with delayed arguments,” *Differential Equations*, vol. 18, pp. 1419–1441, 1983, translated in *Differentsial'nye Uravneniya*, Vol. 18, no. 12, pp. 2027-2050.

[2] A. A. Boichuk and A. M. Samoilenko, *Generalized Inverse Operators and Fredholm Boundary Value Problems*, VSP, Boston, Mass, USA, 2004.

[3] J. Hale, *Theory of Functional Differential Equations*, Applied Mathematical Sciences, Vol. 3, Springer, New York, NY, USA, 2nd edition, 1977.

[4] J. Mallet-Paret, “The fredholm alternative for functional-differential equations of mixed type,” *Journal of Dynamics and Differential Equations*, vol. 11, no. 1, pp. 1–47, 1999.

[5] D. Ya. Khusainov and G. V. Shuklin, “Relative controllability in systems with pure delay,” *International Applied Mechanics*, vol. 41, no. 2, pp. 210–221, 2005, translated in *Prikladnaya Mekhanika*, Vol. 41, no. 2, pp. 118-130.

to Contents, page 1

[6] A. A. Boichuk, J. Diblík, D. Khusainov, and M. Růžičková, "Fredholm's boundary-value problems for differential systems with a single delay," *Nonlinear Analysis, Theory, Methods and Applications*, vol. 72, no. 5, pp. 2251–2258, 2010.

[7] A. A. Boichuk, J. Diblík, D. Y. A. Khusainov, and M. Růžičková, "Boundary value problems for delay differential systems," *Advances in Difference Equations*, vol. 2010, Article ID 593834, 10 pages, 2010.

[8] M. Medvěd, M. Pospíšil, and L. Škripková, "Stability and the nonexistence of blowing-up solutions of nonlinear delay systems with linear parts defined by permutable matrices," *Nonlinear Analysis, Theory, Methods and Applications*, vol. 74, no. 12, pp. 3903–3911, 2011.

[9] A. A. Boichuk and M. K. Grammatikopoulos, "Perturbed fredholm boundary value problems for delay differential systems," *Abstract and Applied Analysis*, no. 15, pp. 843–864, 2003.

[10] A. A. Boichuk, V. F. Zhuravlev, and A. M. Samoilenko, "Generalized inverse operators and noether boundary-value problems," in *Proceedings of the Institute of Mathematics of the National Academy of Sciences of the Ukraine*, vol. 13, p. 320, Kyiv, Ukrania, 1995.

[11] I. G. Malkin, *Some Problems in the Theory of Nonlinear Oscillations*, Gostekhizdat, Moscow, Russia, 1956.

[12] M. A. Krasnosel'skii, G. M. Vainikko, P. P. Zabreiko, Y. A. B. Rutitskii, and V. Ya. Stetsenko, *Approximate Solution of Operator Equations*, Nauka, Moscow, Russia, 1968, Translation: Noordhoff, Groningen, 1972.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 571795, 7 pages
doi:10.1155/2011/571795

Research Article

Bounds of Solutions of Integrodifferential Equations

Zdeněk Šmarda

Department of Mathematics, Faculty of Electrical Engineering and Communication, Technická 8, Brno University of Technology, 61600 Brno, Czech Republic

Correspondence should be addressed to Zdeněk Šmarda, smarda@feec.vutbr.cz

Received 20 January 2011; Accepted 24 February 2011

Academic Editor: Miroslava Růžičková

Some new integral inequalities are given, and bounds of solutions of the following integro-differential equation are determined: $x'(t)-\digamma(t,x(t),\int_0^t k(t,s,x(t),x(s))ds) = h(t)$, $x(0) = x_0$, where $h : R_+ \to R$, $k : R_+^2 \times R^2 \to R$, $\digamma : R_+ \times R^2 \to R$ are continuous functions, $R_+ = [0,\infty)$.

1. Introduction

Ou Yang [1] established and applied the following useful nonlinear integral inequality.

Theorem 1.1. *Let u and h be nonnegative and continuous functions defined on R_+ and let $c \geq 0$ be a constant. Then, the nonlinear integral inequality*

$$u^2(t) \leq c^2 + 2\int_0^t h(s)u(s)ds, \quad t \in R_+ \tag{1.1}$$

implies

$$u(t) \leq c + \int_0^t h(s)ds, \quad t \in R_+. \tag{1.2}$$

This result has been frequently used by authors to obtain global existence, uniqueness, boundedness, and stability of solutions of various nonlinear integral, differential, and

to Contents, page 1

integrodifferential equations. On the other hand, Theorem 1.1 has also been extended and generalized by many authors; see, for example, [2–19]. Like Gronwall-type inequalities, Theorem 1.1 is also used to obtain *a priori* bounds to unknown functions. Therefore, integral inequalities of this type are usually known as Gronwall-Ou Yang type inequalities.

In the last few years there have been a number of papers written on the discrete inequalities of Gronwall inequality and its nonlinear version to the Bihari type, see [13, 16, 20]. Some applications discrete versions of integral inequalities are given in papers [21–23].

Pachpatte [11, 12, 14–16] and Salem [24] have given some new integral inequalities of the Gronwall-Ou Yang type involving functions and their derivatives. Lipovan [7] used the modified Gronwall-Ou Yang inequality with logarithmic factor in the integrand to the study of wave equation with logarithmic nonlinearity. Engler [5] used a slight variant of the Haraux's inequality for determination of global regular solutions of the dynamic antiplane shear problem in nonlinear viscoelasticity. Dragomir [3] applied his inequality to the stability, boundedness, and asymptotic behaviour of solutions of nonlinear Volterra integral equations.

In this paper, we present new integral inequalities which come out from above-mentioned inequalities and extend Pachpatte's results (see [11, 16]) especially. Obtained results are applied to certain classes of integrodifferential equations.

2. Integral Inequalities

Lemma 2.1. *Let u, f, and g be nonnegative continuous functions defined on R_+. If the inequality*

$$u(t) \le u_0 + \int_0^t f(s)\left(u(s) + \int_0^s g(\tau)(u(s) + u(\tau))d\tau\right)ds \tag{2.1}$$

holds where u_0 is a nonnegative constant, $t \in R_+$, then

$$u(t) \le u_0\left[1 + \int_0^t f(s)\exp\left(\int_0^s \left(2g(\tau) + f(\tau)\left(1 + \int_0^\tau g(\sigma)d\sigma\right)\right)d\tau\right)ds\right] \tag{2.2}$$

for $t \in R_+$.

Proof. Define a function $v(t)$ by the right-hand side of (2.1)

$$v(t) = u_0 + \int_0^t f(s)\left(u(s) + \int_0^s g(\tau)(u(s) + u(\tau))d\tau\right)ds. \tag{2.3}$$

Then, $v(0) = u_0$, $u(t) \le v(t)$ and

$$\begin{aligned} v'(t) &= f(t)u(t) + f(t)\int_0^t g(s)(u(t) + u(s))ds \\ &\le f(t)v(t) + f(t)\int_0^t g(s)(v(t) + v(s))ds. \end{aligned} \tag{2.4}$$

to Contents, page 1

Define a function $m(t)$ by

$$m(t) = v(t) + \int_0^t g(s)v(s)ds + v(t)\int_0^t g(s)ds, \tag{2.5}$$

then $m(0) = v(0) = u_0$, $v(t) \le m(t)$,

$$v'(t) \le f(t)m(t), \tag{2.6}$$

$$\begin{aligned} m'(t) &= 2g(t)v(t) + v'(t)\left(1 + \int_0^t g(s)ds\right) \\ &\le m(t)\left[2g(t) + f(t)\left(1 + \int_0^t g(s)ds\right)\right]. \end{aligned} \tag{2.7}$$

Integrating (2.7) from 0 to t, we have

$$m(t) \le u_0 \exp\left(\int_0^t \left(2g(s) + f(s)\left(1 + \int_0^s g(\sigma)d\sigma\right)\right)ds\right). \tag{2.8}$$

Using (2.8) in (2.6), we obtain

$$v'(t) \le u_0 f(t) \exp\left(\int_0^t \left(2g(s) + f(s)\left(1 + \int_0^s g(\sigma)d\sigma\right)\right)ds\right). \tag{2.9}$$

Integrating from 0 to t and using $u(t) \le v(t)$, we get inequality (2.2). The proof is complete. □

Lemma 2.2. *Let u, f, and g be nonnegative continuous functions defined on R_+, $w(t)$ be a positive nondecreasing continuous function defined on R_+. If the inequality*

$$u(t) \le w(t) + \int_0^t f(s)\left(u(s) + \int_0^s g(\tau)(u(s) + u(\tau))d\tau\right)ds, \tag{2.10}$$

holds, where u_0 is a nonnegative constant, $t \in R_+$, then

$$u(t) \le w(t)\left[1 + \int_0^t f(s)\exp\left(\int_0^s \left(2g(\tau) + f(\tau)\left(1 + \int_0^\tau g(\sigma)d\sigma\right)\right)d\tau\right)ds\right], \tag{2.11}$$

where $t \in R_+$.

to Contents, page 1

Proof. Since the function $w(t)$ is positive and nondecreasing, we obtain from (2.10)

$$\frac{u(t)}{w(t)} \leq 1 + \int_0^t f(s)\left(\frac{u(s)}{w(s)} + \int_0^s g(\tau)\left(\frac{u(s)}{w(s)} + \frac{u(\tau)}{w(\tau)}\right)d\tau\right)ds. \tag{2.12}$$

Applying Lemma 2.1 to inequality (2.12), we obtain desired inequality (2.11). □

Lemma 2.3. *Let u, f, g, and h be nonnegative continuous functions defined on R_+, and let c be a nonnegative constant.*

If the inequality

$$u^2(t) \leq c^2 + 2\left[\int_0^t f(s)u(s)\left(u(s) + \int_0^s g(\tau)(u(\tau) + u(s))d\tau\right) + h(s)u(s)\right]ds \tag{2.13}$$

holds for $t \in R_+$, then

$$u(t) \leq p(t)\left[1 + \int_0^t f(s)\exp\left(\int_0^s \left(2g(\tau) + f(\tau)\left(1 + \int_0^\tau g(\sigma)d\sigma\right)\right)d\tau\right)ds\right], \tag{2.14}$$

where

$$p(t) = c + \int_0^t h(s)ds. \tag{2.15}$$

Proof. Define a function $z(t)$ by the right-hand side of (2.13)

$$z(t) = c^2 + 2\left[\int_0^t f(s)u(s)\left(u(s) + \int_0^s g(\tau)(u(\tau) + u(s))d\tau\right) + h(s)u(s)\right]ds. \tag{2.16}$$

Then $z(0) = c^2$, $u(t) \leq \sqrt{z(t)}$ and

$$\begin{aligned} z'(t) &= 2\left[f(t)u(t)\left(u(t) + \int_0^t g(s)(u(t) + u(s))ds\right) + h(t)u(t)\right] \\ &\leq 2\sqrt{z(t)}\left[f(t)\left(\sqrt{z(t)} + \int_0^t g(s)\left(\sqrt{z(t)} + \sqrt{z(s)}\right)ds\right) + h(t)\right]. \end{aligned} \tag{2.17}$$

Differentiating $\sqrt{z(t)}$ and using (2.17), we get

$$\begin{aligned} \frac{d}{dt}\left(\sqrt{z(t)}\right) &= \frac{z'(t)}{2\sqrt{z(t)}} \\ &\leq f(t)\left(\sqrt{z(t)} + \int_0^t g(s)\left(\sqrt{z(t)} + \sqrt{z(s)}\right)ds\right) + h(t). \end{aligned} \tag{2.18}$$

to Contents, page 1

Integrating inequality (2.18) from 0 to t, we have

$$\sqrt{z(t)} \leq p(t) + \int_0^t f(s)\left(\sqrt{z(s)} + \int_0^s g(\tau)\left(\sqrt{z(s)} + \sqrt{z(\tau)}\right)d\tau\right)ds, \tag{2.19}$$

where $p(t)$ is defined by (2.15), $p(t)$ is positive and nondecreasing for $t \in R_+$. Now, applying Lemma 2.2 to inequality (2.19), we get

$$\sqrt{z(t)} \leq p(t)\left[1 + \int_0^t f(s)\exp\left(\int_0^s \left(2g(\tau) + f(\tau)\left(1 + \int_0^\tau g(\sigma)d\sigma\right)\right)d\tau\right)ds\right]. \tag{2.20}$$

Using (2.20) and the fact that $u(t) \leq \sqrt{z(t)}$, we obtain desired inequality (2.14). □

3. Application of Integral Inequalities

Consider the following initial value problem

$$x'(t) - \mathcal{F}\left(t, x(t), \int_0^t k(t, s, x(t), x(s))ds\right) = h(t), \quad x(0) = x_0, \tag{3.1}$$

where $h : R_+ \to R$, $k : R_+^2 \times R^2 \to R$, $\mathcal{F} : R_+ \times R^2 \to R$ are continuous functions. We assume that a solution $x(t)$ of (3.1) exists on R_+.

Theorem 3.1. *Suppose that*

$$\begin{gathered} |k(t, s, u_1, u_2)| \leq f(t)g(s)(|u_1| + |u_2|) \quad \textit{for } (t, s, u_1, u_2) \in R_+^2 \times R^2, \\ |\mathcal{F}(t, u_1, v_1)| \leq f(t)|u_1| + |v_1| \quad \textit{for } (t, u_1, v_1) \in R_+ \times R^2, \end{gathered} \tag{3.2}$$

where f, g are nonnegative continuous functions defined on R_+. Then, for the solution $x(t)$ of (3.1) the inequality

$$\begin{gathered} |x(t)| \leq r(t)\left[1 + \int_0^t f(s)\exp\left(\int_0^s \left(2g(\tau) + f(\tau)\left(1 + \int_0^\tau g(\sigma)d\sigma\right)\right)d\tau\right)ds\right], \\ r(t) = |x_0| + \int_0^t |h(t)|dt \end{gathered} \tag{3.3}$$

holds on R_+.

Proof. Multiplying both sides of (3.1) by $x(t)$ and integrating from 0 to t we obtain

$$x^2(t) = x_0^2 + 2\int_0^t \left[x(s)\mathcal{F}\left(s, x(s), \int_0^s k(s, \tau, x(s), x(\tau))d\tau\right) + x(s)h(s)\right]ds. \tag{3.4}$$

to Contents, page 1

From (3.2) and (3.4), we get

$$|x(t)|^2 \leq |x_0|^2 + 2\int_0^t \left[f(s)|x(s)| \times \left(|x(s)| + \int_0^s g(\tau)(|x(s)| + |x(\tau)|)d\tau \right) + |h(s)||x(s)| \right] ds. \tag{3.5}$$

Using inequality (2.14) in Lemma 2.3, we have

$$|x(t)| \leq r(t)\left[1 + \int_0^t f(s) \exp\left(\int_0^s \left(2g(\tau) + f(\tau)\left(1 + \int_0^\tau g(\sigma)d\sigma \right) \right) d\tau \right) ds \right], \tag{3.6}$$

where

$$r(t) = |x_0| + \int_0^t |h(t)|dt, \tag{3.7}$$

which is the desired inequality (3.3). □

Remark 3.2. It is obvious that inequality (3.3) gives the bound of the solution $x(t)$ of (3.1) in terms of the known functions.

Acknowledgment

This author was supported by the Council of Czech Government grant MSM 00216 30503 and MSM 00216 30529 and by the Grant FEKTS-11-2-921 of Faculty of Electrical Engineering and Communication.

References

[1] L. Ou Yang, "The boundedness of solutions of linear differential equations $y' + A(t)y = 0$," *Advances in Mathematics*, vol. 3, pp. 409–415, 1957.
[2] D. Baınov and P. Simeonov, *Integral Inequalities and Applications*, vol. 57 of *Mathematics and Its Applications (East European Series)*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1992.
[3] S. S. Dragomir, "On Volterra integral equations with kernels of L-type," *Analele Universit a tii din Timi soara. Seria Stiin te Matematice*, vol. 25, no. 2, pp. 21–41, 1987.
[4] S. S. Dragomir and Y.-H. Kim, "On certain new integral inequalities and their applications," *JIPAM: Journal of Inequalities in Pure and Applied Mathematics*, vol. 3, no. 4, article 65, p. 8, 2002.
[5] H. Engler, "Global regular solutions for the dynamic antiplane shear problem in nonlinear viscoelasticity," *Mathematische Zeitschrift*, vol. 202, no. 2, pp. 251–259, 1989.
[6] A. Haraux, *Nonlinear Evolution Equations. Global Behavior of Solutions*, vol. 841 of *Lecture Notes in Mathematics*, Springer, Berlin, Germany, 1981.
[7] O. Lipovan, "A retarded integral inequality and its applications," *Journal of Mathematical Analysis and Applications*, vol. 285, no. 2, pp. 436–443, 2003.
[8] Q. H. Ma and L. Debnath, "A more generalized Gronwall-like integral inequality wit applications," *International Journal of Mathematics and Mathematical Sciences*, vol. 15, pp. 927–934, 2003.
[9] Q.-H. Ma and E.-H. Yang, "On some new nonlinear delay integral inequalities," *Journal of Mathematical Analysis and Applications*, vol. 252, no. 2, pp. 864–878, 2000.
[10] F. W. Meng and W. N. Li, "On some new integral inequalities and their applications," *Applied Mathematics and Computation*, vol. 148, no. 2, pp. 381–392, 2004.

to Contents, page 1

[11] B. G. Pachpatte, "On some new inequalities related to certain inequalities in the theory of differential equations," *Journal of Mathematical Analysis and Applications*, vol. 189, no. 1, pp. 128–144, 1995.

[12] B. G. Pachpatte, "On some integral inequalities similar to Bellman-Bihari inequalities," *Journal of Mathematical Analysis and Applications*, vol. 49, pp. 794–802, 1975.

[13] B. G. Pachpatte, "On certain nonlinear integral inequalities and their discrete analogues," *Facta Universitatis. Series: Mathematics and Informatics*, no. 8, pp. 21–34, 1993.

[14] B. G. Pachpatte, "On some fundamental integral inequalities arising in the theory of differential equations," *Chinese Journal of Contemporary Mathematics*, vol. 22, pp. 261–273, 1994.

[15] B. G. Pachpatte, "On a new inequality suggested by the study of certain epidemic models," *Journal of Mathematical Analysis and Applications*, vol. 195, no. 3, pp. 638–644, 1995.

[16] B. G. Pachpatte, *Inequalities for Differential and Integral Equations*, Mathematics in Science and Engineering 197, Academic Press, San Diego, Calif, USA, 2006.

[17] Z. Šmarda, "Generalization of certain integral inequalities," in *Proceedings of the 8th International Conference on Applied Mathematics (APLIMAT '09)*, pp. 223–228, Bratislava, Slovakia, 2009.

[18] E. H. Yang, "On asymptotic behaviour of certain integro-differential equations," *Proceedings of the American Mathematical Society*, vol. 90, no. 2, pp. 271–276, 1984.

[19] C.-J. Chen, W.-S. Cheung, and D. Zhao, "Gronwall-Bellman-type integral inequalities and applications to BVPs," *Journal of Inequalities and Applications*, vol. 2009, Article ID 258569, 15 pages, 2009.

[20] E. H. Yang, "On some new discrete generalizations of Gronwall's inequality," *Journal of Mathematical Analysis and Applications*, vol. 129, no. 2, pp. 505–516, 1988.

[21] J. Baštinec and J. Diblík, "Asymptotic formulae for a particular solution of linear nonhomogeneous discrete equations," *Computers & Mathematics with Applications*, vol. 45, no. 6–9, pp. 1163–1169, 2003.

[22] J. Diblík, E. Schmeidel, and M. Růžičková, "Existence of asymptotically periodic solutions of system of Volterra difference equations," *Journal of Difference Equations and Applications*, vol. 15, no. 11-12, pp. 1165–1177, 2009.

[23] J. Diblík, E. Schmeidel, and M. Růžičková, "Asymptotically periodic solutions of Volterra system of difference equations," *Computers & Mathematics with Applications*, vol. 59, no. 8, pp. 2854–2867, 2010.

[24] S. Salem, "On some systems of two discrete inequalities of gronwall type," *Journal of Mathematical Analysis and Applications*, vol. 208, no. 2, pp. 553–566, 1997.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 743815, 15 pages
doi:10.1155/2011/743815

Research Article

Compatible and Incompatible Nonuniqueness Conditions for the Classical Cauchy Problem

Josef Diblík[1,2] and Christine Nowak[3]

[1] *Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, University of Technology, 602 00 Brno, Czech Republic*
[2] *Department of Mathematics, Faculty of Electrical Engineering and Communication, Brno University of Technology, 616 00 Brno, Czech Republic*
[3] *Institute for Mathematics, University of Klagenfurt, 9020 Klagenfurt, Austria*

Correspondence should be addressed to Christine Nowak, christine.nowak@uni-klu.ac.at

Received 21 September 2010; Revised 23 November 2010; Accepted 10 March 2011

Academic Editor: Yuri V. Rogovchenko

In the first part of this paper sufficient conditions for nonuniqueness of the classical Cauchy problem $\dot{x} = f(t, x)$, $x(t_0) = x_0$ are given. As the essential tool serves a method which estimates the "distance" between two solutions with an appropriate Lyapunov function and permits to show that under certain conditions the "distance" between two different solutions vanishes at the initial point. In the second part attention is paid to conditions that are obtained by a formal inversion of uniqueness theorems of Kamke-type but cannot guarantee nonuniqueness because they are incompatible.

1. Introduction

Consider the initial value problem

$$\dot{x} = f(t, x), \qquad x(t_0) = x_0, \tag{1.1}$$

where $t_0 \in \mathbb{R}$, $t \in J := [t_0, t_0 + a]$ with $a > 0$, $x, x_0 \in \mathbb{R}^n$ and $f : J \times \mathbb{R}^n \to \mathbb{R}^n$.

In the first part (Section 2) we give sufficient conditions for nonuniqueness of the classical n-dimensional Cauchy problem (1.1). As the essential tool serves a method which estimates the "distance" between two solutions with an appropriate Lyapunov function and permits to show that under certain conditions the "distance" between two different solutions vanishes at the initial point. In the second part (Section 3) we analyze for the one-dimensional case a set of conditions that takes its origin in an inversion of the uniqueness theorem by Kamke (see, e.g., [1, page 56]) but cannot guarantee nonuniqueness since it contains an

to Contents, page 1

inner contradiction. Several attempts were made to get nonuniqueness criteria by using conditions that are (in a certain sense) reverse uniqueness conditions of Kamke type. But this inversion process has to be handled very carefully. It can yield incompatible conditions. This is illustrated by a general set of conditions (in Theorems 3.2, 3.5 and 3.6) that would ensure nonuniqueness, but unfortunately they are inconsistent.

In this paper we study Cauchy problems where f is continuous at the initial point. Related results can be found in [1–5]. In literature there are several investigations for the discontinuous case [1, 6–13] with different qualitative behaviour.

2. Main Result

In the following let $\mathbb{R}_+ := [0, \infty)$, $b > 0$, $\rho > 0$ and

$$S^n_\rho(x_0) := \{x \in \mathbb{R}^n : \|x - x_0\| < \rho\}, \tag{2.1}$$

where $\|\cdot\|$ means the Euclidean norm.

Definition 2.1. We say that the initial value problem (1.1) has at least two different solutions on the interval J if there exist solutions $\varphi(t)$, $\psi(t)$ defined on J and $\varphi \not\equiv \psi$.

The following notions are used in our paper (see, e.g., [14, pages 136 and 137]).

Definition 2.2. A function $\varphi : [0, \rho) \to \mathbb{R}_+$ is said to belong to the class $\mathcal{K}_\rho$ if it is continuous, strictly increasing on $[0, \rho)$ and $\varphi(0) = 0$.

Definition 2.3. A function $V : J \times S^n_\rho(0) \to \mathbb{R}_+$ with $V(t, 0) \equiv 0$ is said to be positive definite if there exists a function $\varphi \in \mathcal{K}_\rho$ such that the relation

$$V(t, x) \geq \varphi(\|x\|) \tag{2.2}$$

is satisfied for $(t, x) \in J \times S^n_\rho(0)$.

For the convenience of the reader we recall the definition of a uniformly Lipschitzian function with respect to a given variable.

Definition 2.4. A function $V(t, \cdot) : S^n_\rho(0) \to \mathbb{R}_+$ is said to be Lipschitzian uniformly with respect to $t \in J$ if for arbitrarily given $x^* \in S^n_\rho(0)$ there exists a constant $k = k(x^*)$ such that

$$\|V(t, x_1^*) - V(t, x_2^*)\| \leq k\|x_1^* - x_2^*\| \tag{2.3}$$

holds for every $t \in J$ and for every x_1^*, x_2^* within a small neighbourhood of x^* in $S^n_\rho(0)$.

In [1, 15, 16] generalized derivatives of a Lipschitzian function along solutions of an associated differential system are analyzed. A slight modification of Theorem 4.3 [15, Appendix I] is the following lemma.

Lemma 2.5. *Let* $V : J \times S^n_\rho(0) \to \mathbb{R}_+$ *be continuous and let* $V(t, \cdot) : S^n_\rho(0) \to \mathbb{R}_+$ *be Lipschitzian uniformly with respect to* $t \in J$. *Let* $x_1, x_2 : J \to S^n_\rho(0)$ *be any two solutions of*

$$\dot{x} = f(t, x), \tag{2.4}$$

where $f : J \times \mathbb{R}^n \to \mathbb{R}^n$ *is a continuous function. Then for the upper right Dini derivative the equality*

$$\begin{aligned} &D^+V(t, x_2(t) - x_1(t)) \\ &\quad := \limsup_{h \to 0^+} \frac{1}{h}\left[V(t+h, x_2(t+h) - x_1(t+h)) - V(t, x_2(t) - x_1(t))\right] \\ &\quad = \limsup_{h \to 0^+} \frac{1}{h}\left[V\big(t+h, x_2(t) - x_1(t) + h\big(f(t, x_2(t)) - f(t, x_1(t))\big)\big) - V(t, x_2(t) - x_1(t))\right] \end{aligned} \tag{2.5}$$

holds.

In the proof of Theorem 2.8 we require the following lemmas which are slight adaptations of Theorem 1.4.1 [14, page 15] and Theorem 1.3.1 [1, page 10] for the left side of the initial point.

Lemma 2.6. *Let* E *be an open* (t, u)*-set in* $\mathbb{R}^2$, *let* $g : E \to \mathbb{R}$ *be a continuous function, and let* u *be the unique solution of*

$$\dot{u} = g(t, u), \quad u(t_2) = u_2, \tag{2.6}$$

to the left with $t_2 > t_0$, $(t_2, u_2) \in E$. *Further, we assume that the scalar continuous function* $m : (t_0, t_2] \to \mathbb{R}$ *with* $(t, m(t)) \in E$ *satisfies* $m(t_2) \le u(t_2)$ *and*

$$D^+m(t) \ge g(t, m(t)), \quad t_0 < t \le t_2. \tag{2.7}$$

Then

$$m(t) \le u(t) \tag{2.8}$$

holds as far as the solution u *exists left of* t_2 *in* $(t_0, t_2]$.

Lemma 2.7. *Let* $S := \{(t, x) : t_0 - a \le t \le t_0,\ |x - x_0| \le b\}$ *and* $f : S \to \mathbb{R}$ *be continuous and nondecreasing in* x *for each fixed* t *in* $[t_0 - a, t_0]$. *Then, the initial value problem* (1.1) *has at most one solution in* $[t_0 - a, t_0]$.

Theorem 2.8 (main result). *Suppose that*

(i) $f : J \times S^n_b(x_0) \to \mathbb{R}^n$ *is a continuous function such that*

$$M := \sup\{\|f(t, x)\| : t \in J,\ x \in S^n_b(x_0)\} < \frac{b}{a}. \tag{2.9}$$

to Contents, page 1

Let x_1 *be a solution of problem* (1.1) *on* J*. Let, moreover, there exist numbers* $t_1 \in (t_0, t_0+a]$, $r \in (0, 2b)$ *and continuous functions* $g : (t_0, t_1] \times \mathbb{R}_+ \to \mathbb{R}$, $V : [t_0, t_1] \times S_r^n(0) \to \mathbb{R}_+$ *such that*

(ii) g *is nondecreasing in the second variable, and the problem*

$$\dot{u} = g(t, u), \qquad \lim_{t \to t_0+} u(t) = 0 \tag{2.10}$$

has a positive solution u^* *on* $(t_0, t_1]$*;*

(iii) V *is positive definite and* $V(t, \cdot) : S_r^n(0) \to \mathbb{R}_+$ *is Lipschitzian uniformly with respect to* $t \in J$*;*

(iv) *for* $t_0 < t \le t_1$, $\|y - x_1(t)\| < r$, *the inequality*

$$\dot{V}(t, y - x_1(t)) \ge g(t, V(t, y - x_1(t))) \tag{2.11}$$

holds where

$$\begin{aligned} &\dot{V}(t, y - x_1(t)) \\ &\quad := \limsup_{h \to 0^+} \frac{1}{h} \left[V(t + h, y - x_1(t) + h[f(t, y) - f(t, x_1(t))]) - V(t, y - x_1(t)) \right]. \end{aligned} \tag{2.12}$$

Then the set of different solutions of problem (1.1) *on interval* J *has the cardinality of the continuum.*

Remark 2.9. If condition (i) is fulfilled then, as it is well known, problem (1.1) is globally solvable and every global solution admits the estimate

$$\|x(t) - x_0\| \le M(t - t_0), \quad t \in J. \tag{2.13}$$

Moreover, for any local solution x_* of problem (1.1), defined on some interval $[t_0, t_1] \subset J$, there exists a global solution x of that problem such that $x(t) = x_*(t)$ for $t \in [t_0, t_1]$.

Remark 2.10. For the case $M = 0$ the initial value problem is unique and the assumptions of Theorem 2.8 cannot be satisfied. Therefore, without loss of generality, we assume $M > 0$ in the proof below.

Proof. At first we show that (1.1) has at least two different solutions on $[t_0, t_1^*]$, where $t_1^* \le t_1$, $t_1^* \le t_0 + \min\{a, b/(3M)\}$ is sufficiently close to t_0. We construct a further solution of (1.1) by finding a point (t_2, x_2) not lying on the solution $x_1(t)$ and starting from this point backwards to the initial point (t_0, x_0).

First we show that there exist values t_2 and x_2, $t_0 < t_2 \le t_1^*$, $\|x_2 - x_0\| \le 2b/3$ such that

$$u^*(t_2) = V(t_2, x_2 - x_1(t_2)) \tag{2.14}$$

to Contents, page 1

holds for the nontrivial solution $u^*(t)$ of $\dot{u} = g(t,u)$. From Lemma 2.7 it follows that $u^*(t)$ is determined uniquely to the left by the initial data $(t_2, u^*(t_2))$. We consider the ε-tubes

$$S(\varepsilon) := \left\{(t,x) : t_0 \leq t \leq t_1^*,\ \|x - x_1(t)\| = \varepsilon\right\} \tag{2.15}$$

for $\varepsilon > 0$ around the solution $x_1(t)$. There exists $\varepsilon_1 > 0$ such that $S(\varepsilon)$ with $0 < \varepsilon \leq \varepsilon_1 < r$ is contained in the set

$$\left\{(t,x) : t_0 \leq t \leq t_1^*,\ \|x - x_0\| \leq \frac{2b}{3}\right\}. \tag{2.16}$$

For $0 \leq \delta \leq \varepsilon_1$, $t \in [t_0, t_1^*]$ we define

$$\begin{gathered}\Psi(\delta,t) := \max_{\|x-x_1(t)\|=\delta} V(t, x - x_1(t)),\\ \Psi(\delta) := \max_{t\in[t_0,t_1^*]} \Psi(\delta,t) \equiv \max_{(t,x)\in S(\delta)} V(t, x - x_1(t)).\end{gathered} \tag{2.17}$$

The function $\Psi(\delta,t)$ is continuous in t for $t_0 \leq t \leq t_1^*$. Since $\lim_{\delta\to 0}\Psi(\delta) = 0$, there exists a δ_2, $0 < \delta_2 \leq \min\{\varepsilon_1, b/3\}$, such that $\Psi(\delta_2) \leq u^*(t_1^*)$. It is clear that inequalities

$$\Psi\left(\delta_2, t_1^*\right) \leq \Psi(\delta_2) \leq u^*\left(t_1^*\right) \tag{2.18}$$

and (due to positive definiteness of V)

$$\Psi(\delta_2, t_0) > 0 = \lim_{t\to t_0+} u^*(t) \tag{2.19}$$

hold. We define a function

$$\omega(t) := \Psi(\delta_2, t) - u^*(t), \tag{2.20}$$

continuous on $[t_0, t_1^*]$. Taking into account inequalities $\omega(t_0) > 0$ and $\omega(t_1^*) \leq 0$ we conclude that there exists t_2, $t_0 < t_2 \leq t_1^*$, with

$$\Psi(\delta_2, t_2) = u^*(t_2). \tag{2.21}$$

The value $\Psi(\delta_2, t_2)$ is taken by $V(t_2, x - x_1(t_2))$ at a point $x = x_2$ such that $\|x_2 - x_1(t_2)\| = \delta_2$ and clearly (in view of the construction) $x_2 \neq x_1(t_2)$. The above statement is proved and (2.14) is valid for (t_2, x_2) determined above.

Now consider the initial value problem

$$\dot{x} = f(t,x), \qquad x(t_2) = x_2. \tag{2.22}$$

to Contents, page 1

Obviously $t_2 - t_0 \leq b/(3M)$ since

$$0 < t_2 - t_0 \leq t_1^* - t_0 \leq \min\left\{a, \frac{b}{3M}\right\} \leq \frac{b}{3M} \tag{2.23}$$

and $\|x_2 - x_0\| \leq 2b/3$ because

$$\begin{aligned}
\|x_2 - x_0\| &= \|x_2 - x_1(t_2) + x_1(t_2) - x_0\| \\
&\leq \|x_2 - x_1(t_2)\| + \|x_1(t_2) - x_0\| = \delta_2 + \left\|\int_{t_0}^{t_2} f(s, x_1(s))ds\right\| \\
&\leq \delta_2 + M(t_2 - t_0) \leq \delta_2 + M\frac{b}{3M} = \delta_2 + \frac{b}{3} \leq \frac{2b}{3}.
\end{aligned} \tag{2.24}$$

Peano's theorem implies that there exists a solution $x_2(t)$ of problem (2.22) on $t_0 \leq t \leq t_2$. We will show that $x_2(t_0) = x_0$. Set

$$m(t) := V(t, x_2(t) - x_1(t)). \tag{2.25}$$

Note that $m(t_2) = u^*(t_2)$. Lemma 2.5 and condition (iv) imply

$$\begin{aligned}
D^+m(t) &:= \limsup_{h\to 0^+} \frac{m(t+h) - m(t)}{h} \\
&= D^+V(t, x_2(t) - x_1(t)) \\
&= \dot{V}(t, x_2(t) - x_1(t)) \geq g(t, V(t, x_2(t) - x_1(t))) = g(t, m(t))
\end{aligned} \tag{2.26}$$

for $t_0 < t \leq t_2$.

Applying Lemma 2.6 we get $m(t) \leq u^*(t)$ for $t_0 < t \leq t_2$. As $m(t) \geq 0$ for $t_0 < t \leq t_2$ and m is continuous at t_0, we find $m(t_0) = 0$. Therefore we have $x_2(t_0) = x_1(t_0) = x_0$ and, as noted above, $x_2(t_2) = x_2 \neq x_1(t_2)$. Thus problem (1.1) has two different solutions.

According to the well-known Kneser theorem [17, Theorem 4.1, page 15] the set of solutions of problem (1.1) either consists of one element or has the cardinality of the continuum. Consequently, if problem (1.1) has two different solutions on interval $[t_0, t_1^*]$ and condition (i) is satisfied, then the set of different solutions of problem (1.1) on interval J has the cardinality of the continuum. The proof is completed. □

Remark 2.11. Note that in the scalar case with $V(t,x) := |x|$ condition (2.11) has the form

$$\big(f(t,y) - f(t, x_1(t))\big) \cdot \operatorname{sign}\big(y - x_1(t)\big) \geq g\big(t, \big|y - x_1(t)\big|\big). \tag{2.27}$$

to Contents, page 1

Example 2.12. Consider for $a = 0.1$, $b = 1$, $t_0 = 0$ and $x_0 = 0$ the scalar differential equation

$$\dot{x} = f(t,x) := \begin{cases} 2x^{1/3} - \dfrac{1}{2} \cdot t^{1/2} \cdot \sin \dfrac{|x|}{t} & \text{if } t \neq 0, \\ 2x^{1/3} & \text{if } t = 0, \end{cases} \tag{2.28}$$

with the initial condition $x(0) = 0$. Let us show that the set of different solutions of this problem on interval J has the cardinality of $\mathbb{R}$. Obviously we can set $x_1(t) \equiv 0$. Put

$$g(t,u) := 2u^{1/3} - \frac{1}{2} \cdot t^{1/2}, \qquad u^*(t) := t^{3/2}, \qquad V(t,x) := |x|. \tag{2.29}$$

Conditions (i), (ii), and (iii) are satisfied. Let us verify that the last condition (iv) is valid, too. We get

$$\begin{aligned} \dot{V}(t, y - x_1(t)) = \dot{V}(t,y) &= (\operatorname{sign} y) \cdot \left[2y^{1/3} - \frac{1}{2} \cdot t^{1/2} \cdot \sin \frac{|y|}{t} \right] \\ &\geq 2|y|^{1/3} - \frac{1}{2} \cdot t^{1/2} = 2V(t,y)^{1/3} - \frac{1}{2} \cdot t^{1/2} = g(t, V(t,y)) \\ &= g(t, V(t, y - x_1(t))). \end{aligned} \tag{2.30}$$

Thus, all conditions of Theorem 2.8 hold and, consequently, the set of different solutions on J of given problem has the cardinality of $\mathbb{R}$.

3. Incompatible Conditions

In this section we show that the formulation of condition (iv) in Theorem 2.8 without knowledge of a solution of the Cauchy problem can lead to an incompatible set of conditions. In the proof of Theorem 3.2 for the one-dimensional case we use the following result given by Nekvinda [18, page 1].

Lemma 3.1. *Let $D \subset \mathbb{R}^2$ and let $f : D \to \mathbb{R}$ be a continuous function in D. Let equation*

$$\dot{x} = f(t,x) \tag{3.1}$$

has the property of left uniqueness. For any $t_0 \in \mathbb{R}$ let A be the set of all $x_0 \in \mathbb{R}$ such that $(t_0, x_0) \in D$ and, for some $\varepsilon > 0$, the initial-value problem (1.1) has more than one solution in the interval $[t_0, t_0+\varepsilon)$. Then A is at most countable.

Theorem 3.2. *The set of conditions (i)–(iv):*

(i) *$f : R_0 \to \mathbb{R}$ with $R_0 := \{(t,x) \in J \times \mathbb{R},\ |x - x_0| \leq b\}$ is continuous;*

to Contents, page 1

(ii) $g : (t_0, t_0 + a] \times (0, \infty) \to \mathbb{R}_+$ *is continuous, nondecreasing in the second variable, and has the following property: there exists a continuous function* $u^*(t)$ *on* J*, which satisfies the differential equation*

$$\dot{u}(t) = g(t, u) \tag{3.2}$$

for $t_0 < t \leq t_0 + a$ *with* $u^*(t_0) = 0$ *and does not vanish for* $t \neq t_0$;

(iii) $V : J \times S^1_{2b}(0) \to \mathbb{R}_+$ *is continuous, positive definite, and Lipschitzian uniformly with respect to* $t \in J$;

(iv) *for* $t_0 < t \leq t_0 + a$, $|x - x_0| \leq b$, $|y - x_0| \leq b$, $x \neq y$,

$$\dot{V}(t, x - y) \geq g(t, V(t, x - y)), \tag{3.3}$$

where we define

$$\dot{V}(t, x - y) := \limsup_{h \to 0^+} \frac{1}{h}\left[V(t + h, x - y + h[f(t, x) - f(t, y)]) - V(t, x - y)\right] \tag{3.4}$$

contains a contradiction.

Proof. Any initial value problem

$$\dot{x} = f(t, x), \qquad x(t_0) = x^* \tag{3.5}$$

with $|x^* - x_0| \leq b$ has at least two different solutions due to Theorem 2.8. Thus we have an uncountable set of nonuniqueness points. We show that solutions passing through different initial points are left unique. Suppose that it does not hold. Let $x_1(t)$ be a solution starting from (t_0, x_1), and let $x_2(t)$ be a solution starting from (t_0, x_2) with $x_2 \neq x_1$. If we assume that these solutions cross at a point $t_1 > t_0$ and if we set

$$m(t) := V(t, x_1(t) - x_2(t)) \tag{3.6}$$

then $m(t_0) > 0$, $m(t_1) = 0$. Therefore there exists a point $t \in (t_0, t_1)$ such that (we apply Lemma 2.5)

$$D^+m(t) = D^+V(t, x_1(t) - x_2(t)) = \dot{V}(t, x_1(t) - x_2(t)) < 0, \tag{3.7}$$

in contradiction to (3.3). Thus we obtain left uniqueness. From Lemma 3.1 we conclude in contrast to the above conclusion that the set of nonuniqueness points (t_0, x^*) can be at most countable. □

In [1, Theorem 1.24.1, page 99] the following nonuniqueness result (see [14, Theorem 2.2.7, page 55], too) is given which uses an inverse Kamke's condition (condition (3.9) below).

to Contents, page 1

Theorem 3.3. *Let $g(t,u)$ be continuous on $0 < t \leq a$, $0 \leq u \leq 2b$, $g(t,0) \equiv 0$, and $g(t,u) > 0$ for $u > 0$. Suppose that, for each t_1, $0 < t_1 < a$, $u(t) \not\equiv 0$ is a differentiable function on $0 < t < t_1$, and continuous on $0 \leq t < t_1$ for which $\dot{u}_+(0)$ exists,*

$$\begin{gathered} \dot{u} = g(t,u), \quad 0 < t < t_1, \\ u(0) = \dot{u}_+(0) = 0. \end{gathered} \tag{3.8}$$

Let $f \in C[R_0, \mathbb{R}]$, where $R_0 : 0 \leq t \leq a$, $|x| \leq b$, and, for $(t,x), (t,y) \in R_0$, $t \neq 0$,

$$\left| f(t,x) - f(t,y) \right| \geq g(t, |x-y|). \tag{3.9}$$

Then, the scalar problem $\dot{x} = f(t,x)$, $x(0) = 0$ has at least two solutions on $0 \leq t \leq a$.

Remark 3.4. In the proof of Theorem 3.3 at first $f(t,0) = 0$ is assumed. Putting $y = 0$ in (3.9) leads to the inequality

$$\left| f(t,x) \right| \geq g(t, |x|). \tag{3.10}$$

As $f(t,x)$ is continuous and $g(t,u) > 0$ for $u > 0$ it follows that $f(t,x)$ must have constant sign for each of the half planes $x > 0$ and $x < 0$. For the upper half plane this implies that

$$\begin{gathered} f(t,x) \geq g(t,x), \\ f(t,x) \leq -g(t,x). \end{gathered} \tag{3.11}$$

For the first inequality nonuniqueness is shown in [1]. But a similar argumentation cannot be used for the second inequality as the following example in [5] shows. We consider the initial value problem $\dot{x} = f(t,x)$, $x(0) = 0$, with

$$f(t,x) = \begin{cases} -\sqrt{x} & \text{if } x \geq 0 \\ \sqrt{-x} & \text{if } x < 0 \end{cases} \tag{3.12}$$

and $g(t,u) := \sqrt{u}$. Thus inequality $|f(t,x)| = \sqrt{|x|} \geq g(t,|x|)$ holds. In the upper half-plane we have $f(t,x) \leq -g(t,x)$. The function $u(t) = t^2/4$ is a nontrivial solution of the comparison equation. Therefore all assumptions are fulfilled, but the initial value problem has at most one solution because of Theorem 1.3.1 [1, page 10].

The next theorem analyzes in the scalar case (for $(t_0, x_0) = (0,0)$) that even fulfilling a rather general condition (see condition (3.14) in the following theorem) cannot ensure nonuniqueness since the set of all conditions contains an inner contradiction. The proof was motivated by the paper [5].

to Contents, page 1

Theorem 3.5. *There exists no system of three functions f, g, and V satisfying the following suppositions:*

(i) $f : R_0 \to \mathbb{R}$ *with* $R_0 := \{(t,x) \in \mathbb{R} \times \mathbb{R},\ 0 \le t \le a, 0 \le x \le b\}$ *is a continuous function;*

(ii) *the continuous function* $g : (0,a] \times \mathbb{R}_+ \to \mathbb{R}_+$, $g(t,0) := 0$ *if* $t \in (0,a]$, *has the following property: there exists a continuously differentiable function* $u^*(t)$ *on* $0 \le t \le a$, *satisfying the differential equation*

$$\dot{u} = g(t,u) \tag{3.13}$$

for $0 < t \le a$ *such that* $u^*(0) = 0$ *and* $u^*(t) > 0$ *for* $t \ne 0$;

(iii) *the continuous function* $V : [0,a] \times S_b^1(0) \to \mathbb{R}_+$ *is positive definite, and for all* $0 < t \le a$, $0 < x < b$ *continuously differentiable;*

(iv) *for* $0 < t \le a$, $0 < y < x \le b$,

$$\dot{V}(t, x-y) \ge g(t, V(t, x-y)) \ge 0, \tag{3.14}$$

where we define

$$\dot{V}(t,x-y) := V_1'(t,x-y) + V_2'(t,x-y) \cdot [f(t,x) - f(t,y)] \tag{3.15}$$

and subscript indices denote the derivative with respect to the first and second argument, respectively;

(v) *there exist a positive constant* ϑ *and a function* $\xi : (0,b] \to (0,\infty)$ *such that for* $0 < t \le a$ *and* $0 < x \le b$

$$\begin{gathered} 0 \le V'_1(t,x) \le \vartheta \cdot \xi(x), \quad 0 < V_2'(t,x) \le \vartheta \cdot \frac{\xi(x)}{x}, \\ V(t,x) \ge \xi(x); \end{gathered} \tag{3.16}$$

(vi) *for* $t \in [0,a]$ *and* x, y *with* $0 < y < x \le b$ *the inequality*

$$f(t,x) - f(t,y) \ge 0 \tag{3.17}$$

holds.

Proof. Let us show that the above properties are not compatible. For fixed numbers x, y with $0 < y < x \le b$ consider the auxiliary function

$$F(t) := \frac{f(t,x) - f(t,y)}{x-y} + 1, \quad t \in [0,a]\ . \tag{3.18}$$

to Contents, page 1

Clearly, F is continuous and assumes a (positive) maximum. Set

$$K = \max_{[0,a]} F(t) \geq 1. \tag{3.19}$$

If the function g fulfills the inequality

$$g(t,u) \leq \Lambda \cdot u \tag{3.20}$$

with a positive constant Λ in a domain $0 < t \leq A \leq a$, $0 \leq u \leq B$, $B > 0$, then the initial value problem

$$\dot{u} = g(t,u), \qquad u(0) = 0 \tag{3.21}$$

has the unique trivial solution $u = 0$. Really, since $u^*(t) > 0$ for $t \in (0,a]$, by integrating inequality

$$\frac{\dot{u}^*(t)}{u^*(t)} \leq \Lambda \tag{3.22}$$

with limits t, $A^* \in (0,A)$ we get

$$u^*(A^*) \leq u^*(t) \exp[\Lambda(A^* - t)] \tag{3.23}$$

and for $t \longrightarrow 0^+$

$$u^*(A^*) \leq 0 \tag{3.24}$$

which contradicts positivity of u^*. Therefore problem (3.21) has only the trivial solution. Hence, there exist a sequence $\{(t_n,u_n)\}$ with $t_n \in (0,a]$, $u_n > 0$, $\lim_{n\to\infty}(t_n,u_n) = (0,0)$ and a sequence $\{\lambda_n\}$, $\lambda_n > 0$, with $\lim_{n\to\infty}\lambda_n = \infty$ such that the inequality

$$g(t_n,u_n) > \lambda_n u_n \tag{3.25}$$

holds for every n. Consider now the relation

$$V(t,x) = 0. \tag{3.26}$$

Due to the properties of V we conclude that for all sufficiently small positive numbers t_n, u_n (i.e., for all sufficiently large n) there exists a (sufficiently small and positive) number $\tilde{u}_n$ such that the equation

$$V(t_n,x) = u_n \tag{3.27}$$

to Contents, page 1

has the solution $x = \tilde{u}_n$. Thus a sequence $\{\tilde{u}_n\}$ with $\tilde{u}_n > 0$ and $\lim_{n\to\infty}\tilde{u}_n = 0$ corresponds to the sequence $\{(t_n, u_n)\}$. For every n define a number j_n as

$$j_n = \left\lceil \frac{x-y}{\tilde{u}_n} - 1 \right\rceil, \tag{3.28}$$

where $\lceil\cdot\rceil$ is the ceiling function. Without loss of generality we can suppose that

$$\frac{x-y}{\tilde{u}_n} > 4. \tag{3.29}$$

Obviously,

$$\frac{x-y}{\tilde{u}_n} - 1 \le j_n < \frac{x-y}{\tilde{u}_n}. \tag{3.30}$$

Moreover, without loss of generality we can suppose that for every sufficiently large n the inequality

$$\lambda_n > 2\vartheta K \tag{3.31}$$

holds. Set

$$\begin{aligned} x_0 &:= y, \\ x_1 &:= y + \tilde{u}_n, \\ x_2 &:= y + 2\tilde{u}_n, \\ &\ \vdots \\ x_{j_n} &:= y + j_n \cdot \tilde{u}_n, \\ x_{j_n+1} &:= x. \end{aligned} \tag{3.32}$$

Consider for all sufficiently large n the expression

$$\mathcal{E}_n := j_n V_1'(t_n, \tilde{u}_n) + V_2'(t_n,\ \tilde{u}_n) \cdot \left[f(t_n, x) - f(t_n, y)\right]. \tag{3.33}$$

to Contents, page 1

Then

$$
\begin{aligned}
\mathcal{E}_n &= j_n V_1'(t_n,\tilde{u}_n) + V_2'(t_n,\tilde{u}_n)\cdot\sum_{i=1}^{j_n+1}\left[f(t_n,x_i)-f(t_n,x_{i-1})\right]\\
&= j_n V_1'(t_n,\tilde{u}_n) + V_2'(t_n,\tilde{u}_n)\cdot\sum_{i=1}^{j_n}\left[f(t_n,x_i)-f(t_n,x_{i-1})\right] + \left[f(t_n,x)-f(t_n,x_{j_n})\right]\\
&\geq [\text{due to (vi)}] \geq j_n V_1'(t_n,\tilde{u}_n) + V_2'(t_n,\tilde{u}_n)\cdot\sum_{i=1}^{j_n}\left[f(t_n,x_i)-f(t_n,x_{i-1})\right]\\
&= [\text{due to (iv) and (v)}] = j_n V_1'(t_n,\tilde{u}_n) + V_2'(t_n,\tilde{u}_n)\cdot j_n\left[\frac{-V_1'(t_n,\tilde{u}_n)+\dot{V}(t_n,\tilde{u}_n)}{V_2'(t_n,\tilde{u}_n)}\right]\\
&\geq [\text{due to (iv)}]\\
&\geq j_n V_1'(t_n,\tilde{u}_n) + V_2'(t_n,\tilde{u}_n)\cdot j_n\left[\frac{-V_1'(t_n,\tilde{u}_n)+g(t_n,V(t_n,\tilde{u}_n))}{V'_2(t_n,\tilde{u}_n)}\right]\\
&= j_n\cdot g(t_n,V(t_n,\tilde{u}_n)) = [\text{due to (3.27)}] = j_n\cdot g(t_n,u_n) \geq [\text{due to (3.25)}]\\
&\geq j_n\lambda_n u_n \geq [\text{due to (3.31)}] \geq j_n u_n\cdot 2\vartheta K \geq [\text{due to (3.30)}]\\
&\geq \left(\frac{x-y}{\tilde{u}_n}-1\right) u_n\cdot 2\vartheta K\\
&= \left(\frac{x-y}{\tilde{u}_n}-1\right)V(t_n,\tilde{u}_n)\cdot 2\vartheta K\\
&\geq [\text{due to (3.16) }] \geq \left(\frac{x-y}{\tilde{u}_n}-1\right)\xi(\tilde{u}_n)\cdot 2\vartheta K\\
&= (x-y-\tilde{u}_n)\cdot\frac{\xi(\tilde{u}_n)}{\tilde{u}_n}\cdot 2\vartheta K\\
&\geq [\text{due to (3.29)}] \geq \frac{3}{4}\cdot(x-y)\cdot 2\vartheta K\cdot\frac{\xi(\tilde{u}_n)}{\tilde{u}_n}\\
&= \frac{3}{2}\cdot(x-y)\cdot\vartheta K\cdot\frac{\xi(\tilde{u}_n)}{\tilde{u}_n} > 0.
\end{aligned}
\tag{3.34}
$$

Estimating the expression $\mathcal{E}_n$ from above we get (see (3.32))

$$
\begin{aligned}
\mathcal{E}_n &\leq \frac{x-y}{\tilde{u}_n}V_1'(t_n,\tilde{u}_n) + V_2'(t_n,\tilde{u}_n)\cdot\left[f(t_n,x)-f(t_n,y)\right]\\
&\leq [\text{due to (v)}]\\
&\leq \frac{x-y}{\tilde{u}_n}\cdot\vartheta\cdot\xi(\tilde{u}_n) + \vartheta\cdot\frac{\xi(\tilde{u}_n)}{\tilde{u}_n}\cdot(K-1)(x-y) = \vartheta\cdot\frac{\xi(\tilde{u}_n)}{\tilde{u}_n}\cdot K(x-y).
\end{aligned}
\tag{3.35}
$$

to Contents, page 1

These two above estimations yield

$$0 < \frac{3}{2} \cdot (x-y) \cdot \vartheta K \cdot \frac{\xi(\tilde{u}_n)}{\tilde{u}_n} \leq \mathcal{E}_n \leq (x-y) \cdot \vartheta K \cdot \frac{\xi(\tilde{u}_n)}{\tilde{u}_n}, \tag{3.36}$$

in contrast to $(3/2) \nleq 1$. Since the initially taken points x and y, $0 < y < x$, can be chosen arbitrarily close to zero, the theorem is proved. □

The following result is a consequence of Theorem 3.5 if $V(t,x) := |x|$, $\xi(x) := x$ and $\vartheta = 1$. Condition (3.38) below was discussed previously in [5].

Theorem 3.6. *There exists no system of two functions f and g satisfying the following suppositions:*

(i) *$f : R_0 \to \mathbb{R}$ with $R_0 := \{(t,x) \in \mathbb{R} \times \mathbb{R},\ 0 \leq t \leq a, 0 \leq x \leq b\}$ is a continuous function;*

(ii) *the continuous function $g : (0,a] \times \mathbb{R}_+ \to \mathbb{R}_+$, $g(t,0) := 0$ if $t \in (0,a]$, has the following property: there exists a continuously differentiable function $u^*(t)$ on $0 \leq t \leq a$, satisfying the differential equation*

$$\dot{u}(t) = g(t,u) \tag{3.37}$$

for $0 < t \leq a$ such that $u^*(0) = 0$ and $u^*(t) > 0$ for $t \neq 0$;

(iii) *for $0 < t \leq a$, $0 < y < x \leq b$*

$$f(t,x) - f(t,y) \geq g(t,x-y) \geq 0; \tag{3.38}$$

(iv) *for $0 < y < x \leq b$ the inequality $f(0,x) - f(0,y) \geq 0$ holds.*

Remark 3.7. Let us note that in the singular case, that is, when we permit that the function $f(t,x)$ is not continuous at $t = 0$, the given sets of conditions in Theorems 3.5 and 3.6 can be compatible. This can be seen from the proof where the continuity of f is substantial. Such singular case was considered in [13].

Acknowledgments

This research was supported by the Grants P201/11/0768 and P201/10/1032 of the Czech Grant Agency (Prague), by the Project FEKT-S-11-2(921) and by the Council of Czech Government MSM 00216 30529.

References

[1] R. P. Agarwal and V. Lakshmikantham, *Uniqueness and Nonuniqueness Criteria for Ordinary Differential Equations*, vol. 6 of *Series in Real Analysis*, World Scientific Publishing, River Edge, NJ, USA, 1993.
[2] J. Kalas, "General nonuniqueness theorem for ordinary differential equations," *Dynamics of Continuous, Discrete and Impulsive Systems*, vol. 3, no. 1, pp. 97–111, 1997.
[3] J. Kalas, "Nonuniqueness for the solutions of ordinary differential equations," *Czechoslovak Mathematical Journal*, vol. 29, no. 1, pp. 105–112, 1979.

to Contents, page 1

[4] J. Kalas, "Nonuniqueness results for ordinary differential equations," *Czechoslovak Mathematical Journal*, vol. 48, no. 2, pp. 373–384, 1998.

[5] C. Nowak and H. Stettner, "Nonuniqueness results for ordinary differential equations," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 30, no. 6, pp. 3935–3938, 1997.

[6] K. Balla, "Solution of singular boundary value problems for non-linear systems of ordinary differential equations," *U.S.S.R. Computational Mathematics and Mathematical Physics*, vol. 20, no. 4, pp. 100–115, 1980.

[7] V. A. Čečik, "Investigation of systems of ordinary differential equations with a singularity," *Trudy Moskovskogo Matematičeskogo Obščestva*, vol. 8, pp. 155–198, 1959 (Russian).

[8] J. Diblík and C. Nowak, "A nonuniqueness criterion for a singular system of two ordinary differential equations," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 64, no. 4, pp. 637–656, 2006.

[9] J. Diblík and M. Růžičková, "Inequalities for solutions of singular initial problems for Carathéodory systems via Ważewski's principle," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 69, no. 12, pp. 4482–4495, 2008.

[10] J. Diblík and M. Růžičková, "Existence of positive solutions of a singular initial problem for a nonlinear system of differential equations," *The Rocky Mountain Journal of Mathematics*, vol. 34, no. 3, pp. 923–944, 2004.

[11] J. Kalas, "Nonuniqueness theorem for a singular Cauchy problem," *Georgian Mathematical Journal*, vol. 7, no. 2, pp. 317–327, 2000.

[12] I. T. Kiguradze, *Some Singular Boundary Value Problems for Ordinary Differential Equations*, Tbilisi University Press, Tbilisi, Georgia, 1975.

[13] C. Nowak, "Some remarks on a paper by M. Samimi on: "Nonuniqueness criteria for ordinary differential equations"," *Applicable Analysis*, vol. 47, no. 1, pp. 39–44, 1992.

[14] V. Lakshmikantham and S. Leela, *Differential and Integral Inequalities: Theory and Applications. Vol. I: Ordinary Differential Equations*, Academic Press, New York, NY, USA, 1969.

[15] T. Rouche, P. Habets, and M. Laloy, *Stability Theory by Liapunov's Direct Method*, Springer, New York, NY, USA, 1977.

[16] T. Yoshizawa, *Stability Theory by Liapunov's Second Method*, The Mathematical Society of Japan, Tokyo, Japan, 1966.

[17] P. Hartman, *Ordinary Differential Equations*, vol. 38 of *Classics in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 2002.

[18] M. Nekvinda, "On uniqueness of solutions of differential equations," *Časopis Pro Pěstování Matematiky*, vol. 108, no. 1, pp. 1–7, 1983.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 814962, 16 pages
doi:10.1155/2011/814962

Research Article

Conjugacy of Self-Adjoint Difference Equations of Even Order

Petr Hasil

Department of Mathematics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic

Correspondence should be addressed to Petr Hasil, hasil@mendelu.cz

Received 31 January 2011; Revised 5 April 2011; Accepted 18 May 2011

Academic Editor: Elena Braverman

We study oscillation properties of $2n$-order Sturm-Liouville difference equations. For these equations, we show a conjugacy criterion using the p-criticality (the existence of linear dependent recessive solutions at ∞ and $-\infty$). We also show the equivalent condition of p-criticality for one term $2n$-order equations.

1. Introduction

In this paper, we deal with $2n$-order Sturm-Liouville difference equations and operators

$$L(y)_k = \sum_{\nu=0}^{n} (-\Delta)^{\nu} \left(r_k^{[\nu]} \Delta^{\nu} y_{k+n-\nu} \right) = 0, \quad r_k^{[n]} > 0,\ k \in \mathbb{Z}, \tag{1.1}$$

where Δ is the forward difference operator, that is, $\Delta y_k = y_{k+1} - y_k$, and $r^{[\nu]}$, $\nu = 0, \dots, n$, are real-valued sequences. The main result is the conjugacy criterion which we formulate for the equation $L(y)_k + q_k y_{k+n} = 0$, that is viewed as a perturbation of (1.1), and we suppose that (1.1) is at least p-critical for some $p \in \{1, \dots, n\}$. The concept of p-criticality (a disconjugate equation is said to be p-critical if and only if it possesses p solutions that are recessive both at ∞ and $-\infty$, see Section 3) was introduced for second-order difference equations in [1], and later in [2] for (1.1). For the continuous counterpart of the used techniques, see [3–5] from where we get an inspiration for our research.

The paper is organized as follows. In Section 2, we recall necessary preliminaries. In Section 3, we recall the concept of p-criticality, as introduced in [2], and show the first

to Contents, page 1

result—the equivalent condition of p-criticality for the one term difference equation

$$\Delta^n \left(r_k \Delta^n y_k \right) = 0 \tag{1.2}$$

(Theorem 3.4). In Section 4 we show the conjugacy criterion for equation

$$(-\Delta)^n \left(r_k \Delta^n y_k \right) + q_k y_{k+n} = 0, \tag{1.3}$$

and Section 5 is devoted to the generalization of this criterion to the equation with the middle terms

$$\sum_{\nu=0}^{n} (-\Delta)^\nu \left(r_k^{[\nu]} \Delta^\nu y_{k+n-\nu} \right) + q_k y_{k+n} = 0. \tag{1.4}$$

2. Preliminaries

The proof of our main result is based on equivalency of (1.1) and the linear Hamiltonian difference systems

$$\Delta x_k = A x_{k+1} + B_k u_k, \qquad \Delta u_k = C_k x_{k+1} - A^T u_k, \tag{2.1}$$

where A, B_k, and C_k are $n \times n$ matrices of which B_k and C_k are symmetric. Therefore, we start this section recalling the properties of (2.1), which we will need later. For more details, see the papers [6–11] and the books [12, 13].

The substitution

$$x_k^{[y]} = \begin{pmatrix} y_{k+n-1} \\ \Delta y_{k+n-2} \\ \vdots \\ \Delta^{n-1} y_k \end{pmatrix}, \qquad u_k^{[y]} = \begin{pmatrix} \sum_{\nu=1}^{n} (-\Delta)^{\nu-1} \left(r_k^{[\nu]} \Delta^\nu y_{k+n-\nu} \right) \\ \vdots \\ -\Delta\left(r_k^{[n]} \Delta^n y_k \right) + r_k^{[n-1]} \Delta^{n-1} y_{k+1} \\ r_k^{[n]} \Delta^n y_k \end{pmatrix} \tag{2.2}$$

transforms (1.1) to linear Hamiltonian system (2.1) with the $n \times n$ matrices A, B_k, and C_k given by

$$\begin{gathered} A = \left(a_{ij} \right)_{i,j=1}^n, \qquad a_{ij} = \begin{cases} 1, & \text{if } j = i+1,\ i = 1, \dots, n-1, \\ 0, & \text{elsewhere}, \end{cases} \\ B_k = \operatorname{diag}\left\{ 0, \dots, 0, \frac{1}{r_k^{[n]}} \right\}, \qquad C_k = \operatorname{diag}\left\{ r_k^{[0]}, \dots, r_k^{[n-1]} \right\}. \end{gathered} \tag{2.3}$$

Then, we say that the solution (x, u) of (2.1) is generated by the solution y of (1.1).

to Contents, page 1

Let us consider, together with system (2.1), the matrix linear Hamiltonian system

$$\Delta X_k = AX_{k+1} + B_k U_k, \qquad \Delta U_k = C_k X_{k+1} - A^T U_k, \tag{2.4}$$

where the matrices A, B_k, and C_k are also given by (2.3). We say that the matrix solution (X, U) of (2.4) is generated by the solutions $y^{[1]}, \dots, y^{[n]}$ of (1.1) if and only if its columns are generated by $y^{[1]}, \dots, y^{[n]}$, respectively, that is, $(X, U) = (x^{[y_1]}, \dots, x^{[y_n]}, u^{[y_1]}, \dots, u^{[y_n]})$. Reversely, if we have the solution (X, U) of (2.4), the elements from the first line of the matrix X are exactly the solutions $y^{[1]}, \dots, y^{[n]}$ of (1.1). Now, we can define the oscillatory properties of (1.1) via the corresponding properties of the associated Hamiltonian system (2.1) with matrices A, B_k, and C_k given by (2.3), for example, (1.1) is disconjugate if and only if the associated system (2.1) is disconjugate, the system of solutions $y^{[1]}, \dots, y^{[n]}$ is said to be recessive if and only if it generates the recessive solution X of (2.4), and so forth. Therefore, we define the following properties just for linear Hamiltonian systems.

For system (2.4), we have an analog of the continuous *Wronskian identity*. Let (X, U) and $(\tilde{X}, \tilde{U})$ be two solutions of (2.4). Then,

$$X_k^T \tilde{U}_k - U_k^T \tilde{X}_k \equiv W \tag{2.5}$$

holds with a constant matrix W. We say that the solution (X, U) of (2.4) is a *conjoined basis*, if

$$X_k^T U_k \equiv U_k^T X_k, \qquad \operatorname{rank}\begin{pmatrix} X \\ U \end{pmatrix} = n. \tag{2.6}$$

Two conjoined bases $(X, U), (\tilde{X}, \tilde{U})$ of (2.4) are called *normalized* conjoined bases of (2.4) if $W = I$ in (2.5) (where I denotes the identity operator).

System (2.1) is said to be *right disconjugate* in a discrete interval $[l, m]$, $l, m \in \mathbb{Z}$, if the solution $\binom{X}{U}$ of (2.4) given by the initial condition $X_l = 0, U_l = I$ satisfies

$$\ker X_{k+1} \subseteq \ker X_k, \qquad X_k X_{k+1}^{\dagger} (I - A)^{-1} B_k \geq 0, \tag{2.7}$$

for $k = l, \dots, m-1$, see [6]. Here ker, †, and $\geq$ stand for kernel, Moore-Penrose generalized inverse, and nonnegative definiteness of the matrix indicated, respectively. Similarly, (2.1) is said to be *left disconjugate* on $[l, m]$, if the solution given by the initial condition $X_m = 0$, $U_m = -I$ satisfies

$$\ker X_k \subseteq \ker X_{k+1}, \quad X_{k+1} X_k^{\dagger} B_k (I - A)^{T-1} \geq 0, \quad k = l, \dots, m-1. \tag{2.8}$$

System (2.1) is disconjugate on $\mathbb{Z}$, if it is right disconjugate, which is the same as left disconjugate, see [14, Theorem 1], on $[l, m]$ for every $l, m \in \mathbb{Z}$, $l < m$. System (2.1) is said to be *nonoscillatory* at ∞ (*nonoscillatory* at $-\infty$), if there exists $l \in \mathbb{Z}$ such that it is right disconjugate on $[l, m]$ for every $m > l$ (there exists $m \in \mathbb{Z}$ such that (2.1) is left disconjugate on $[l, m]$ for every $l < m$).

to Contents, page 1

We call a conjoined basis $\binom{\tilde{X}}{\tilde{U}}$ of (2.4) the *recessive solution* at ∞, if the matrices $\tilde{X}_k$ are nonsingular, $\tilde{X}_k \tilde{X}_{k+1}^{-1}(I-A_k)^{-1}B_k \geq 0$ (both for large k), and for any other conjoined basis $\binom{X}{U}$, for which the (constant) matrix $X^T\tilde{U} - U^T\tilde{X}$ is nonsingular, we have

$$\lim_{k\to\infty} X_k^{-1}\tilde{X}_k = 0. \tag{2.9}$$

The solution (X, U) is called the *dominant solution* at ∞. The recessive solution at ∞ is determined uniquely up to a right multiple by a nonsingular constant matrix and exists whenever (2.4) is nonoscillatory and eventually controllable. (System is said to be *eventually controllable* if there exist $N, \kappa \in \mathbb{N}$ such that for any $m \geq N$ the trivial solution $\binom{x}{u} = \binom{0}{0}$ of (2.1) is the only solution for which $x_m = x_{m+1} = \cdots = x_{m+\kappa} = 0$.) The equivalent characterization of the recessive solution $\binom{\tilde{X}}{\tilde{U}}$ of eventually controllable Hamiltonian difference systems (2.1) is

$$\lim_{k\to\infty} \left(\sum k\tilde{X}_{j+1}^{-1}(I-A)^{-1}B_j\tilde{X}_j^{T-1}\right)^{-1} = 0, \tag{2.10}$$

see [12]. Similarly, we can introduce the recessive and the dominant solutions at $-\infty$. For related notions and results for second-order dynamic equations, see, for example, [15, 16].

We say that a pair (x, u) is *admissible* for system (2.1) if and only if the first equation in (2.1) holds.

The energy functional of (1.1) is given by

$$\mathcal{F}(y) := \sum_{k=-\infty}^{\infty}\sum_{\nu=0}^{n} r_k^{[\nu]}\left(\Delta^\nu y_{k+n-\nu}\right)^2. \tag{2.11}$$

Then, for admissible (x, u), we have

$$\begin{aligned}\mathcal{F}(y) &= \sum_{k=-\infty}^{\infty}\sum_{\nu=0}^{n} r_k^{[\nu]}\left(\Delta^\nu y_{k+n-\nu}\right)^2 \\ &= \sum_{k=-\infty}^{\infty}\left[\sum_{\nu=0}^{n-1} r_k^{[\nu]}\left(\Delta^\nu y_{k+n-\nu}\right)^2 + \frac{1}{r_k^{[n]}}\left(r_k^{[n]}\Delta^n y_k\right)^2\right] \\ &= \sum_{k=-\infty}^{\infty}\left[x_{k+1}^T C_k x_{k+1} + u_k^T B_k u_k\right] =: \mathcal{F}(x,u).\end{aligned} \tag{2.12}$$

To prove our main result, we use a variational approach, that is, the equivalency of disconjugacy of (1.1) and positivity of $\mathcal{F}(x, u)$; see [6].

Now, we formulate some auxiliary results, which are used in the proofs of Theorems 3.4 and 4.1. The following Lemma describes the structure of the solution space of

$$\Delta^n\left(r_k\Delta^n y_k\right) = 0, \quad r_k > 0. \tag{2.13}$$

to Contents, page 1

Lemma 2.1 (see [17, Section 2]). *Equation* (2.13) *is disconjugate on* $\mathbb{Z}$ *and possesses a system of solutions* $y^{[j]}, \tilde{y}^{[j]}$, $j = 1, \ldots, n$, *such that*

$$y^{[1]} \prec \cdots \prec y^{[n]} \prec \tilde{y}^{[1]} \prec \cdots \prec \tilde{y}^{[n]} \tag{2.14}$$

as $k \to \infty$, *where* $f \prec g$ *as* $k \to \infty$ *for a pair of sequences* f, g *means that* $\lim_{k\to\infty}(f_k/g_k) = 0$. *If* (2.14) *holds, the solutions* $y^{[j]}$ *form the recessive system of solutions at* ∞, *while* $\tilde{y}^{[j]}$ *form the dominant system,* $j = 1, \ldots, n$. *The analogous statement holds for the ordered system of solutions as* $k \to -\infty$.

Now, we recall the transformation lemma.

Lemma 2.2 (see [14, Theorem 4]). *Let* $h_k > 0$, $L(y) = \sum_{\nu=0}^{n}(-\Delta)^{\nu}(r_k^{[\nu]}\Delta^{\nu}y_{k+n-\nu})$ *and consider the transformation* $y_k = h_k z_k$. *Then, one has*

$$h_{k+n}L(y) = \sum_{\nu=0}^{n}(-\Delta)^{\nu}\left(R_k^{[\nu]}\Delta^{\nu}z_{k+n-\nu}\right), \tag{2.15}$$

where

$$R_k^{[n]} = h_{k+n}h_k r_k^{[n]}, \qquad R_k^{[0]} = h_{k+n}L(h), \tag{2.16}$$

that is, y *solves* $L(y) = 0$ *if and only if* z *solves the equation*

$$\sum_{\nu=0}^{n}(-\Delta)^{\nu}\left(R_k^{[\nu]}\Delta^{\nu}z_{k+n-\nu}\right) = 0. \tag{2.17}$$

The next lemma is usually called the second mean value theorem of summation calculus.

Lemma 2.3 (see [17, Lemma 3.2]). *Let* $n \in \mathbb{N}$ *and the sequence* a_k *be monotonic for* $k \in [K+n-1, L+n-1]$ *(i.e.,* Δa_k *does not change its sign for* $k \in [K+n-1, L+n-2]$*). Then, for any sequence* b_k *there exist* $n_1, n_2 \in [K, L-1]$ *such that*

$$\begin{gathered}\sum_{j=K}^{L-1}a_{n+j}b_j \le a_{K+n-1}\sum_{i=K}^{n_1-1}b_i + a_{L+n-1}\sum_{i=n_1}^{L-1}b_i,\\ \sum_{j=K}^{L-1}a_{n+j}b_j \ge a_{K+n-1}\sum_{i=K}^{n_2-1}b_i + a_{L+n-1}\sum_{i=n_2}^{L-1}b_i.\end{gathered} \tag{2.18}$$

Now, let us consider the linear difference equation

$$y_{k+n} + a_k^{[n-1]}y_{k+n-1} + \cdots + a_k^{[0]}y_k = 0, \tag{2.19}$$

where $k \ge n_0$ for some $n_0 \in \mathbb{N}$ and $a_k^{[0]} \neq 0$, and let us recall the main ideas of [18] and [19, Chapter IX].

to Contents, page 1

An integer $m > n_0$ is said to be a *generalized zero* of multiplicity k of a nontrivial solution y of (2.19) if $y_{m-1} \neq 0$, $y_m = y_{m+1} = \cdots = y_{m+k-2} = 0$, and $(-1)^k y_{m-1} y_{m+k-1} \geq 0$. Equation (2.19) is said to be eventually disconjugate if there exists $N \in \mathbb{N}$ such that no non-trivial solution of this equation has n or more generalized zeros (counting multiplicity) on $[N, \infty)$.

A system of sequences $u_k^{[1]}, \ldots, u_k^{[n]}$ is said to form the *D-Markov system* of sequences for $k \in [N, \infty)$ if Casoratians

$$C\left(u^{[1]}, \ldots, u^{[j]}\right)_k = \begin{vmatrix} u_k^{[1]} & \cdots & u_k^{[j]} \\ u_{k+1}^{[1]} & \cdots & u_{k+1}^{[j]} \\ \vdots & & \vdots \\ u_{k+j-1}^{[1]} & \cdots & u_{k+j-1}^{[j]} \end{vmatrix}, \quad j = 1, \ldots, n \tag{2.20}$$

are positive on $(N + j, \infty)$.

Lemma 2.4 (see [19, Theorem 9.4.1]). *Equation* (2.19) *is eventually disconjugate if and only if there exist* $N \in \mathbb{N}$ *and solutions* $y^{[1]}, \ldots, y^{[n]}$ *of* (2.19) *which form a D-Markov system of solutions on* (N, ∞). *Moreover, this system can be chosen in such a way that it satisfies the additional condition*

$$\lim_{k \to \infty} \frac{y_k^{[i]}}{y_k^{[i+1]}} = 0, \quad i = 1, \ldots, n-1. \tag{2.21}$$

3. Criticality of One-Term Equation

Suppose that (1.1) is disconjugate on $\mathbb{Z}$ and let $\hat{y}^{[i]}$ and $\tilde{y}^{[i]}$, $i = 1, \ldots, n$, be the recessive systems of solutions of $L(y) = 0$ at $-\infty$ and ∞, respectively. We introduce the linear space

$$\mathcal{H} = \operatorname{Lin}\left\{\hat{y}^{[1]}, \ldots, \hat{y}^{[n]}\right\} \cap \operatorname{Lin}\left\{\tilde{y}^{[1]}, \ldots, \tilde{y}^{[n]}\right\}. \tag{3.1}$$

Definition 3.1 (see [2]). Let (1.1) be disconjugate on $\mathbb{Z}$ and let $\dim \mathcal{H} = p \in \{1, \ldots, n\}$. Then, we say that the operator L (or (1.1)) is *p-critical* on $\mathbb{Z}$. If $\dim \mathcal{H} = 0$, we say that L is *subcritical* on $\mathbb{Z}$. If (1.1) is not disconjugate on $\mathbb{Z}$, that is, $L \ngeq 0$, we say that L is *supercritical* on $\mathbb{Z}$.

To prove the result in this section, we need the following statements, where we use the generalized power function

$$k^{(0)} = 1, \qquad k^{(i)} = k(k-1)\cdots(k-i+1), \quad i \in \mathbb{N}. \tag{3.2}$$

For reader's convenience, the first statement in the following lemma is slightly more general than the corresponding one used in [2] (it can be verified directly or by induction).

to Contents, page 1

Lemma 3.2 (see [2]). *The following statements hold.*

(i) *Let z_k be any sequence, $m \in \{0, \ldots, n\}$, and*

$$y_k := \sum_{j=0}^{k-1} (k-j-1)^{(n-1)} z_j, \tag{3.3}$$

then

$$\Delta^m y_k = \begin{cases} (n-1)^{(m)} \displaystyle\sum_{j=0}^{k-1} (k-j-1)^{(n-1-m)} z_j, & m \le n-1, \\ (n-1)! z_k, & m = n. \end{cases} \tag{3.4}$$

(ii) *The generalized power function has the binomial expansion*

$$(k-j)^{(n)} = \sum_{i=0}^{n} (-1)^i \binom{n}{i} k^{(n-i)} (j+i-1)^{(i)}. \tag{3.5}$$

We distinguish two types of solutions of (2.13). The *polynomial* solutions $k^{(i)}$, $i = 0, \ldots, n-1$, for which $\Delta^n y_k = 0$, and *nonpolynomial* solutions

$$\sum_{j=0}^{k-1} (k-j-1)^{(n-1)} j^{(i)} r_j^{-1}, \quad i = 0, \ldots, n-1, \tag{3.6}$$

for which $\Delta^n y_k \neq 0$. (Using Lemma 3.2(i) we obtain $\Delta^n y_k = (n-1)! k^{(i)} r_k^{-1}$.)

Now, we formulate one of the results of [20].

Proposition 3.3 (see [20, Theorem 4]). *If for some $m \in \{0, \ldots, n-1\}$*

$$\sum_{k=-\infty}^{0} \left[k^{(n-m-1)}\right]^2 r_k^{-1} = \infty = \sum_{k=0}^{\infty} \left[k^{(n-m-1)}\right]^2 r_k^{-1}, \tag{3.7}$$

then

$$\operatorname{Lin}\left\{1, \ldots, k^{(m)}\right\} \subseteq \mathscr{A}, \tag{3.8}$$

that is, (2.13) *is at least* $(m+1)$*-critical on* $\mathbb{Z}$.

Now, we show that (3.7) is also sufficient for (2.13) to be at least $(m+1)$-critical.

Theorem 3.4. *Let $m \in \{0, \ldots, n-1\}$. Equation* (2.13) *is at least* $(m+1)$*-critical if and only if* (3.7) *holds.*

to Contents, page 1

Proof. Let $\mathcal{U}^+$ and $\mathcal{U}^-$ denote the subspaces of the solution space of (2.13) generated by the recessive system of solutions at ∞ and $-\infty$, respectively. Necessity of (3.7) follows directly from Proposition 3.3. To prove sufficiency, it suffices to show that if one of the sums in (3.7) is convergent, then $\{1, \dots, k^{(m)}\} \nsubseteq \mathcal{U}^+ \cap \mathcal{U}^-$. We show this statement for the sum $\sum^\infty$. The other case is proved similarly, so it will be omitted. Particularly, we show

$$\sum_{k=0}^{\infty} \left[k^{(n-m-1)}\right]^2 r_k^{-1} < \infty \Longrightarrow k^{(m)} \notin \mathcal{U}^+. \tag{3.9}$$

Let us denote $p := n - m - 1$, and let us consider the following nonpolynomial solutions of (2.13):

$$y_k^{[\ell]} = \sum_{j=0}^{k-1} (k-j-1)^{(n-1)} j^{(p+\ell-1)} r_j^{-1} - \sum_{i=0}^{p} \left[(-1)^i \binom{n-1}{i} (k-1)^{(n-1-i)} \sum_{j=0}^{\infty} j^{(p+\ell-1)} (j+i-1)^{(i)} r_j^{-1}\right], \tag{3.10}$$

where $\ell = 1-p, \dots, m+1$. By Stolz-Cesàro theorem, since (using Lemma 3.2(i)) $\Delta^n y_k^{[\ell]} = (n-1)! k^{(p+\ell-1)} r_k^{-1}$, these solutions are ordered, that is, $y^{[i]} \prec y^{[i+1]}$, $i = 1-p, \dots, m$, as well as the polynomial solutions, that is, $k^{(i)} \prec k^{(i+1)}$, $i = 0, \dots, n-2$.

By some simple calculation and by Lemma 3.2 (at first, we use (i), and at the end, we use (ii)), we have

$$\begin{aligned}
&\Delta^m y_k^{[1]} \\
&\quad = \frac{(n-1)!}{(n-m-1)!} \sum_{j=0}^{k-1} (k-j-1)^{(n-m-1)} j^{(p)} r_j^{-1} \\
&\qquad - \sum_{i=0}^{p} \left[(-1)^i \binom{n-1}{i} \frac{(n-1-i)!}{(n-m-1-i)!} (k-1)^{(n-m-1-i)} \sum_{j=0}^{\infty} j^{(p)} (j+i-1)^{(i)} r_j^{-1}\right] \\
&\quad = \frac{(n-1)!}{p!} \sum_{j=0}^{k-1} (k-j-1)^{(p)} j^{(p)} r_j^{-1} \\
&\qquad - \sum_{i=0}^{p} \left[(-1)^i \frac{(n-1)!(n-1-i)!}{(n-1-i)! i! (p-i)!} (k-1)^{(p-i)} \sum_{j=0}^{\infty} j^{(p)} (j+i-1)^{(i)} r_j^{-1}\right] \\
&\quad = \frac{(n-1)!}{p!} \left\{ \sum_{j=0}^{k-1} (k-j-1)^{(p)} j^{(p)} r_j^{-1} - \sum_{i=0}^{p} \left[(-1)^i \binom{p}{i} (k-1)^{(p-i)} \sum_{j=0}^{\infty} j^{(p)} (j+i-1)^{(i)} r_j^{-1}\right]\right\} \\
&\quad = \frac{(n-1)!}{p!} \left[\sum_{j=0}^{k-1} (k-j-1)^{(p)} j^{(p)} r_j^{-1} - \sum_{j=0}^{\infty} (k-j-1)^{(p)} j^{(p)} r_j^{-1}\right]
\end{aligned}$$

to Contents, page 1

$$= -\frac{(n-1)!}{p!} \sum_{j=k}^{\infty} (k-j-1)^{(p)} j^{(p)} r_j^{-1}$$

$$= (-1)^{p+1} \frac{(n-1)!}{p!} \sum_{j=k}^{\infty} (j+1-k)^{(p)} j^{(p)} r_j^{-1},$$

$$\sum_{j=k}^{\infty} (j+1-k)^{(p)} j^{(p)} r_j^{-1} \le \sum_{j=k}^{\infty} \left[j^{(p)} \right]^2 r_j^{-1}. \tag{3.11}$$

Hence, from this and by Stolz-Cesàro theorem, we get

$$\lim_{k\to\infty} \frac{y_k^{[1]}}{k^{(m)}} = \frac{1}{m!} \lim_{k\to\infty} \Delta^m y_k^{[1]} = 0, \tag{3.12}$$

thus $y_k^{[1]} \prec k^{(m)}$. We obtained that $\{1, k, \ldots, k^{(m-1)}, y^{[1-p]}, \ldots, y^{[1]}\} \prec k^{(m)}$, which means that we have n solutions less than $k^{(m)}$, therefore $k^{(m)} \notin \mathcal{U}^{+}$ and (2.13) is at most m-critical. □

4. Conjugacy of Two-Term Equation

In this section, we show the conjugacy criterion for two-term equation.

Theorem 4.1. *Let $n > 1$, q_k be a real-valued sequence, and let there exist an integer $m \in \{0, \ldots, n-1\}$ and real constants $c_0, \ldots, c_m$ such that* (2.13) *is at least* $(m+1)$*-critical and the sequence $h_k := c_0 + c_1 k + \cdots + c_m k^{(m)}$ satisfies*

$$\limsup_{K\downarrow -\infty,\, L\uparrow\infty} \sum_{k=K}^{L} q_k h_{k+n}^2 \le 0. \tag{4.1}$$

If $q \not\equiv 0$, then

$$(-\Delta)^n (r_k \Delta^n y_k) + q_k y_{k+n} = 0 \tag{4.2}$$

is conjugate on $\mathbb{Z}$.

Proof. We prove this theorem using the variational principle; that is, we find a sequence $y \in \ell_0^2(\mathbb{Z})$ such that the energy functional $F(y) = \sum_{k=-\infty}^{\infty} [r_k (\Delta^n y_k)^2 + q_k y_{k+n}^2] < 0$.

At first, we estimate the first term of $F(y)$. To do this, we use the fact that this term is an energy functional of (2.13). Let us denote it by $\tilde{F}$ that is,

$$\tilde{F}(y) = \sum_{k=-\infty}^{\infty} r_k (\Delta^n y_k)^2. \tag{4.3}$$

to Contents, page 1

Using the substitution (2.2), we find out that (2.13) is equivalent to the linear Hamiltonian system (2.1) with the matrix $C_k \equiv 0$; that is,

$$\Delta x_k = A_k x_{k+1} + B_k u_k, \qquad \Delta u_k = -A^T u_k, \tag{4.4}$$

and to the matrix system

$$\Delta X_k = A_k X_{k+1} + B_k U_k, \qquad \Delta U_k = -A^T U_k. \tag{4.5}$$

Now, let us denote the recessive solutions of (4.5) at $-\infty$ and ∞ by (X^-, U^-) and (X^+, U^+), respectively, such that the first $m+1$ columns of X^+ and X^- are generated by the sequences $1, k, \ldots, k^{(m)}$. Let K, L, M, and N be arbitrary integers such that $N - M > 2n$, $M - L > 2n$, and $L - K > 2n$ (some additional assumptions on the choice of K, L, M, N will be specified later), and let $(x^{[f]}, u^{[f]})$ and $(x^{[g]}, u^{[g]})$ be the solutions of (4.4) given by the formulas

$$\begin{aligned}
x_k^{[f]} &= X_k^- \left(\sum_{j=K}^{k-1} \mathcal{B}_j^-\right)\left(\sum_{j=K}^{L-1} \mathcal{B}_j^-\right)^{-1} (X_L^-)^{-1} x_L^{[h]},\\
u_k^{[f]} &= U_k^- \left(\sum_{j=K}^{k-1} \mathcal{B}_j^-\right)\left(\sum_{j=K}^{L-1} \mathcal{B}_j^-\right)^{-1} (X_L^-)^{-1} x_L^{[h]} + (X_k^-)^{T-1}\left(\sum_{j=K}^{L-1} \mathcal{B}_j^-\right)^{-1} (X_L^-)^{-1} x_L^{[h]},\\
x_k^{[g]} &= X_k^+ \left(\sum_{j=k}^{N-1} \mathcal{B}_j^+\right)\left(\sum_{j=M}^{N-1} \mathcal{B}_j^+\right)^{-1} (X_M^+)^{-1} x_M^{[h]},\\
u_k^{[g]} &= U_k^+ \left(\sum_{j=k}^{N-1} \mathcal{B}_j^+\right)\left(\sum_{j=M}^{N-1} \mathcal{B}_j^+\right)^{-1} (X_M^+)^{-1} x_M^{[h]} - (X_k^+)^{T-1}\left(\sum_{j=M}^{N-1} \mathcal{B}_j^+\right)^{-1} (X_M^+)^{-1} x_M^{[h]},
\end{aligned} \tag{4.6}$$

where

$$\begin{aligned}
\mathcal{B}_k^- &= (X_{k+1}^-)^{-1}(I-A)^{-1} B_k (X_k^-)^{T-1},\\
\mathcal{B}_k^+ &= (X_{k+1}^+)^{-1}(I-A)^{-1} B_k (X_k^+)^{T-1},
\end{aligned} \tag{4.7}$$

and $(x^{[h]}, u^{[h]})$ is the solution of (4.4) generated by h. By a direct substitution, and using the convention that $\sum_k^{k-1} = 0$, we obtain

$$x_K^{[f]} = 0, \qquad x_L^{[f]} = x_L^{[h]}, \qquad x_M^{[g]} = x_M^{[h]}, \qquad x_N^{[g]} = 0. \tag{4.8}$$

Now, from (4.1), together with the assumption $q \not\equiv 0$, we have that there exist $\tilde{k} \in \mathbb{Z}$ and $\varepsilon > 0$ such that $q_{\tilde{k}} \le -\varepsilon$. Because the numbers K, L, M, and N have been "almost free" so far, we may choose them such that $L < \tilde{k} < M - n - 1$.

to Contents, page 1

Let us introduce the test sequence

$$y_k := \begin{cases} 0, & k \in (-\infty, K-1], \\ f_k, & k \in [K, L-1], \\ h_k(1+D_k), & k \in [L, M-1], \\ g_k, & k \in [M, N-1], \\ 0, & k \in [N, \infty), \end{cases} \tag{4.9}$$

where

$$D_k = \begin{cases} \delta > 0, & k = \tilde{k} + n, \\ 0, & \text{otherwise}. \end{cases} \tag{4.10}$$

To finish the first part of the proof, we use (4.4) to estimate the contribution of the term

$$\widetilde{F}(y) = \sum_{k=-\infty}^{\infty} r_k (\Delta^n y_k)^2 = \sum_{k=-\infty}^{\infty} u_k^{[y]T} B_k u_k^{[y]} = \sum_{k=K}^{N-1} u_k^{[y]T} B_k u_k^{[y]}. \tag{4.11}$$

Using the definition of the test sequence y, we can split $\widetilde{F}$ into three terms. Now, we estimate two of them as follows. Using (4.4), we obtain

$$\begin{aligned} \sum_{k=K}^{L-1} u_k^{[f]T} B_k u_k^{[f]} &= \sum_{k=K}^{L-1} \left[u_k^{[f]T} \left(\Delta x_k^{[f]} - A x_{k+1}^{[f]} \right) \right] = \sum_{k=K}^{L-1} \left[u_k^{[f]T} \Delta x_k^{[f]} - u_k^{[f]T} A x_{k+1}^{[f]} \right] \\ &= \sum_{k=K}^{L-1} \left[\Delta \left(u_k^{[f]T} x_k^{[f]} \right) - \Delta u_k^{[f]T} x_{k+1}^{[f]} - u_k^{[f]T} A x_{k+1}^{[f]} \right] \\ &= \sum_{k=K}^{L-1} \left[\Delta \left(u_k^{[f]T} x_k^{[f]} \right) - x_{k+1}^{[f]T} \left(\Delta u_k^{[f]} + A^T u_k^{[f]} \right) \right] = u_k^{[f]T} x_k^{[f]} \Big|_K^L = x_L^{[f]T} u_L^{[f]} \\ &= x_L^{[h]T} \left[U_L^- (X_L^-)^{-1} x_L^{[h]} + (X_L^-)^{T-1} \left(\sum_{j=K}^{L-1} \mathcal{B}_j^- \right)^{-1} (X_L^-)^{-1} x_L^{[h]} \right] \\ &= x_L^{[h]T} (X_L^-)^{T-1} \left(\sum_{j=K}^{L-1} \mathcal{B}_j^- \right)^{-1} (X_L^-)^{-1} x_L^{[h]} =: \mathcal{G}, \end{aligned} \tag{4.12}$$

to Contents, page 1

where we used the fact that $x_L^{[h]T} U_L^{-} (X_L^{-})^{-1} x_L^{[h]} \equiv 0$ (recall that the last $n-m-1$ entries of $x_L^{[h]}$ are zeros and that the first $m+1$ columns of X^{-} and U^{-} are generated by the solutions $1, \dots, k^{(m)}$). Similarly,

$$\sum_{k=M}^{N-1} u_k^{[g]T} B_k u_k^{[g]} = -x_M^{[g]T} u_M^{[g]} = x_M^{[h]T} (X_M^{+})^{T-1} \left(\sum_{j=M}^{N-1} \mathcal{B}_j^{+} \right)^{-1} (X_M^{+})^{-1} x_M^{[h]} =: \mathscr{H}. \tag{4.13}$$

Using property (2.10) of recessive solutions of the linear Hamiltonian difference systems, we can see that $\mathcal{G} \to 0$ as $K \to -\infty$ and $\mathscr{H} \to 0$ as $N \to \infty$. We postpone the estimation of the middle term of $\widetilde{F}$ to the end of the proof.

To estimate the second term of $F(y)$, we estimate at first its terms

$$\sum_{k=K}^{L-1} q_k f_{k+n}^2, \qquad \sum_{k=M}^{N-1} q_k g_{k+n}^2. \tag{4.14}$$

For this estimation, we use Lemma 2.3. To do this, we have to show the monotonicity of the sequences

$$\begin{aligned} &\frac{f_k}{h_k} \quad \text{for } k \in [K+n-1, L+n-1], \\ &\frac{g_k}{h_k} \quad \text{for } k \in [M+n-1, N+n-1]. \end{aligned} \tag{4.15}$$

Let $x^{[1]}, \dots, x^{[2n]}$ be the ordered system of solutions of (2.13) in the sense of Lemma 2.1. Then, again by Lemma 2.1, there exist real numbers $d_1, \dots, d_n$ such that $h = d_1 x^{[1]} + \cdots + d_n x^{[n]}$. Because $h \not\equiv 0$, at least one coefficient d_i is nonzero. Therefore, we can denote $p := \max\{i \in [1,n] : d_i \neq 0\}$, and we replace the solution $x^{[p]}$ by h. Let us denote this new system again $x^{[1]}, \dots, x^{[2n]}$ and note that this new system has the same properties as the original one.

Following Lemma 2.2, we transform (2.13) via the transformation $y_k = h_k z_k$, into

$$\sum_{\nu=0}^{n} (-\Delta)^{\nu} \left(R_k^{[\nu]} \Delta^{\nu} z_{k+n-\nu} \right) = 0, \tag{4.16}$$

that is,

$$(-\Delta)^n \left(r_k h_k h_{k+n} \Delta^{n-1} w_k \right) + \cdots - \Delta \left(R_k^{[1]} w_{k+n-1} \right) = 0 \tag{4.17}$$

to Contents, page 1

possesses the fundamental system of solutions

$$w^{[1]} = -\Delta\left(\frac{x^{[1]}}{h}\right), \dots, w^{[p-1]} = -\Delta\left(\frac{x^{[p-1]}}{h}\right),$$
$$w^{[p]} = \Delta\left(\frac{x^{[p+1]}}{h}\right), \dots, w^{[2n-1]} = \Delta\left(\frac{x^{[2n]}}{h}\right). \tag{4.18}$$

Now, let us compute the Casoratians

$$C\left(w^{[1]}\right) = w^{[1]} = -\Delta\left(\frac{x^{[1]}}{h}\right) = \frac{C\left(x^{[1]}, h\right)}{h_k h_{k+1}} > 0,$$
$$C\left(w^{[1]}, w^{[2]}\right) = \frac{C\left(x^{[1]}, x^{[2]}, h\right)}{h_k h_{k+1} h_{k+2}} > 0,$$
$$\vdots$$
$$C\left(w^{[1]}, \dots, w^{[2n-1]}\right) = \frac{C\left(x^{[1]}, \dots, x^{[p-1]}, x^{[p+1]}, \dots, x^{[2n]}, h\right)}{h_k \cdots h_{k+2n-1}} > 0. \tag{4.19}$$

Hence, $w^{[1]}, \dots, w^{[2n-1]}$ form the D-Markov system of sequences on $[M, \infty)$, for M sufficiently large. Therefore, by Lemma 2.4, (4.17) is eventually disconjugate; that is, it has at most $2n-2$ generalized zeros (counting multiplicity) on $[M, \infty)$. The sequence $\Delta(g/h)$ is a solution of (4.17), and we have that this sequence has generalized zeros of multiplicity $n-1$ both at M and at N; that is,

$$\Delta\left(\frac{g_{M+i}}{h_{M+i}}\right) = 0 = \Delta\left(\frac{g_{N+i}}{h_{N+i}}\right), \quad i = 0, \dots, n-2. \tag{4.20}$$

Moreover, $g_M/h_M = 1$ and $g_N/h_N = 0$. Hence, $\Delta(g_k/h_k) \le 0, k \in [M, N+n-1]$. We can proceed similarly for the sequence f/h.

Using Lemma 2.3, we have that there exist integers $\xi_1 \in [K, L-1]$ and $\xi_2 \in [M, N-1]$ such that

$$\sum_{k=K}^{L-1} q_k f_{k+n}^2 = \sum_{k=K}^{L-1}\left[q_k h_{k+n}^2\left(\frac{f_{k+n}}{h_{k+n}}\right)^2\right] \le \sum_{k=\xi_1}^{L-1} q_k h_{k+n}^2,$$
$$\sum_{k=M}^{N-1} q_k g_{k+n}^2 = \sum_{k=M}^{N-1}\left[q_k h_{k+n}^2\left(\frac{g_{k+n}}{h_{k+n}}\right)^2\right] \le \sum_{k=M}^{\xi_2-1} q_k h_{k+n}^2. \tag{4.21}$$

to Contents, page 1

Finally, we estimate the remaining term of $F(y)$. By (4.9), we have

$$\begin{aligned}
\sum_{k=L}^{M-1}&\left[r_k(\Delta^n y_k)^2+q_k y_{k+n}^2\right]\\
&=\sum_{k=L}^{M-1}\left\{r_k[\Delta^n h_k+\Delta^n(h_k D_k)]^2+q_k(h_{k+n}+h_{k+n}D_{k+n})^2\right\}\\
&=\sum_{k=L}^{M-1}\left\{r_k[\Delta^n(h_k D_k)]^2+q_k h_{k+n}^2+2q_k h_{k+n}^2 D_{k+n}+q_k h_{k+n}^2 D_{k+n}^2\right\}\\
&=\sum_{k=\tilde{k}}^{\tilde{k}+n}\left\{r_k[\Delta^n(h_k D_k)]^2\right\}+\sum_{k=L}^{M-1}\left[q_k h_{k+n}^2\right]+2q_{\tilde{k}}h_{\tilde{k}+n}^2 D_{\tilde{k}+n}+q_{\tilde{k}}h_{\tilde{k}+n}^2 D_{\tilde{k}+n}^2\\
&=\sum_{k=\tilde{k}}^{\tilde{k}+n}\left\{r_k\left[(-1)^{k-\tilde{k}}\binom{n}{k-\tilde{k}}h_{\tilde{k}+n}\delta\right]^2\right\}+\sum_{k=L}^{M-1}\left[q_k h_{k+n}^2\right]+2\delta q_{\tilde{k}}h_{\tilde{k}+n}^2+\delta^2 q_{\tilde{k}}h_{\tilde{k}+n}^2\\
&\le\delta^2 h_{\tilde{k}+n}^2\sum_{k=\tilde{k}}^{\tilde{k}+n}\left[r_k\binom{n}{k-\tilde{k}}^2\right]+\sum_{k=L}^{M-1}\left[q_k h_{k+n}^2\right]-2\delta\varepsilon h_{\tilde{k}+n}^2-\delta^2\varepsilon h_{\tilde{k}+n}^2\\
&<\delta^2 h_{\tilde{k}+n}^2\sum_{k=\tilde{k}}^{\tilde{k}+n}\left[r_k\binom{n}{k-\tilde{k}}^2\right]+\sum_{k=L}^{M-1}\left[q_k h_{k+n}^2\right]-2\delta\varepsilon h_{\tilde{k}+n}^2.
\end{aligned}\tag{4.22}$$

Altogether, we have

$$\begin{aligned}
F(y)&<\delta^2 h_{\tilde{k}+n}^2\sum_{k=\tilde{k}}^{\tilde{k}+n}\left[r_k\binom{n}{k-\tilde{k}}^2\right]+\sum_{k=L}^{M-1}\left[q_k h_{k+n}^2\right]-2\delta\varepsilon h_{\tilde{k}+n}^2+\mathcal{G}+\mathcal{H}+\sum_{k=\xi_1}^{L-1}q_k h_{k+n}^2+\sum_{k=M}^{\xi_2-1}q_k h_{k+n}^2\\
&=\delta^2 h_{\tilde{k}+n}^2\sum_{k=\tilde{k}}^{\tilde{k}+n}\left[r_k\binom{n}{k-\tilde{k}}^2\right]-2\delta\varepsilon h_{\tilde{k}+n}^2+\mathcal{G}+\mathcal{H}+\sum_{k=\xi_1}^{\xi_2-1}q_k h_{k+n}^2,
\end{aligned}\tag{4.23}$$

where for K sufficiently small is $\mathcal{G}<\delta^2/3$, for N sufficiently large is $\mathcal{H}<\delta^2/3$, and, from (4.1), $\sum_{k=\xi_1}^{\xi_2-1}q_k h_{k+n}^2<\delta^2/3$ for $\xi_1<L$ and $\xi_2>M$. Therefore,

$$\begin{aligned}
F(y)&<\delta^2+\delta^2 h_{\tilde{k}+n}^2\sum_{k=\tilde{k}}^{\tilde{k}+n}\left[r_k\binom{n}{k-\tilde{k}}^2\right]-2\delta\varepsilon h_{\tilde{k}+n}^2\\
&=\delta\left\{\delta\left[1+h_{\tilde{k}+n}^2\sum_{k=\tilde{k}}^{\tilde{k}+n}\left[r_k\binom{n}{k-\tilde{k}}^2\right]\right]-\varepsilon h_{\tilde{k}+n}^2\right\},
\end{aligned}\tag{4.24}$$

which means that $F(y)<0$ for δ sufficiently small, and (4.2) is conjugate on $\mathbb{Z}$. □

to Contents, page 1

5. Equation with the Middle Terms

Under the additional condition $q_k \leq 0$ for large $|k|$, and by combining of the proof of Theorem 4.1 with the proof of [2, Lemma 1], we can establish the following criterion for the full $2n$-order equation.

Theorem 5.1. *Let $n > 1$, q_k be a real-valued sequence, and let there exist an integer $m \in \{0, \ldots, n-1\}$ and real constants $c_0, \ldots, c_m$ such that (1.1) is at least $(m+1)$-critical and the sequence $h_k := c_0 + c_1 k + \cdots + c_m k^{(m)}$ satisfies*

$$\limsup_{K\downarrow-\infty, L\uparrow\infty} \sum_{k=K}^{L} q_k h_{k+n}^2 \leq 0. \tag{5.1}$$

If $q_k \leq 0$ for large $|k|$ and $q \not\equiv 0$, then

$$L(y)_k + q_k y_{k+n} = \sum_{\nu=0}^{n} (-\Delta)^\nu \left(r_k^{[\nu]} \Delta^\nu y_{k+n-\nu} \right) + q_k y_{k+n} = 0 \tag{5.2}$$

is conjugate on $\mathbb{Z}$.

Remark 5.2. Using Theorem 3.4, we can see that the statement of Theorem 4.1 holds if and only if (3.7) holds. Finding a criterion similar to Theorem 3.4 for (1.1) is still an open question.

Remark 5.3. In the view of the matrix operator associated to (1.1) in the sense of [21], we can see that the perturbations in Theorem 4.1 affect the diagonal elements of the associated matrix operator. A description of behavior of (1.1), with regard to perturbations of limited part of the associated matrix operator (but not only of the diagonal elements), is given in [2].

Acknowledgment

The research was supported by the Czech Science Foundation under Grant no. P201/10/1032.

References

[1] F. Gesztesy and Z. Zhao, “Critical and subcritical Jacobi operators defined as Friedrichs extensions,” *Journal of Differential Equations*, vol. 103, no. 1, pp. 68–93, 1993.
[2] O. Došlý and P. Hasil, “Critical higher order Sturm-Liouville difference operators,” to appear in *Journal of Difference Equations and Applications*.
[3] O. Došlý, “Existence of conjugate points for selfadjoint linear differential equations,” *Proceedings of the Royal Society of Edinburgh. Section A*, vol. 113, no. 1-2, pp. 73–85, 1989.
[4] O. Došlý, “Oscillation criteria and the discreteness of the spectrum of selfadjoint, even order, differential operators,” *Proceedings of the Royal Society of Edinburgh. Section A*, vol. 119, no. 3-4, pp. 219–232, 1991.
[5] O. Došlý and J. Komenda, “Conjugacy criteria and principal solutions of self-adjoint differential equations,” *Archivum Mathematicum*, vol. 31, no. 3, pp. 217–238, 1995.
[6] M. Bohner, “Linear Hamiltonian difference systems: disconjugacy and Jacobi-type conditions,” *Journal of Mathematical Analysis and Applications*, vol. 199, no. 3, pp. 804–826, 1996.

to Contents, page 1

[7] M. Bohner, O. Došlý, and W. Kratz, "A Sturmian theorem for recessive solutions of linear Hamiltonian difference systems," *Applied Mathematics Letters*, vol. 12, no. 2, pp. 101–106, 1999.
[8] O. Došlý, "Transformations of linear Hamiltonian difference systems and some of their applications," *Journal of Mathematical Analysis and Applications*, vol. 191, no. 2, pp. 250–265, 1995.
[9] O. Došlý, "Symplectic difference systems: oscillation theory and hyperbolic Prüfer transformation," *Abstract and Applied Analysis*, vol. 2004, no. 4, pp. 285–294, 2004.
[10] L. H. Erbe and P. X. Yan, "Disconjugacy for linear Hamiltonian difference systems," *Journal of Mathematical Analysis and Applications*, vol. 167, no. 2, pp. 355–367, 1992.
[11] L. H. Erbe and P. X. Yan, "Qualitative properties of Hamiltonian difference systems," *Journal of Mathematical Analysis and Applications*, vol. 171, no. 2, pp. 334–345, 1992.
[12] C. D. Ahlbrandt and A. C. Peterson, *Discrete Hamiltonian Systems: Difference Equations, Continued Fractions, and Riccati Equations*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2010.
[13] W. Kratz, *Quadratic Functionals in Variational Analysis and Control Theory*, vol. 6 of *Mathematical Topics*, Akademie, Berlin, Germany, 1995.
[14] M. Bohner and O. Došlý, "Disconjugacy and transformations for symplectic systems," *The Rocky Mountain Journal of Mathematics*, vol. 27, no. 3, pp. 707–743, 1997.
[15] M. Bohner and S. Stević, "Trench's perturbation theorem for dynamic equations," *Discrete Dynamics in Nature and Society*, vol. 2007, Article ID 75672, 11 pages, 2007.
[16] M. Bohner and S. Stević, "Linear perturbations of a nonoscillatory second-order dynamic equation," *Journal of Difference Equations and Applications*, vol. 15, no. 11-12, pp. 1211–1221, 2009.
[17] O. Došlý, "Oscillation criteria for higher order Sturm-Liouville difference equations," *Journal of Difference Equations and Applications*, vol. 4, no. 5, pp. 425–450, 1998.
[18] P. Hartman, "Difference equations: disconjugacy, principal solutions, Green's functions, complete monotonicity," *Transactions of the American Mathematical Society*, vol. 246, pp. 1–30, 1978.
[19] R. P. Agarwal, *Difference Equations and Inequalities, Theory, Methods, and Applications*, vol. 155 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 1992.
[20] P. Hasil, "Criterion of p-criticality for one term $2n$-order difference operators," *Archivum Mathematicum*, vol. 47, pp. 99–109, 2011.
[21] W. Kratz, "Banded matrices and difference equations," *Linear Algebra and Its Applications*, vol. 337, pp. 1–20, 2001.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 565067, 21 pages
doi:10.1155/2011/565067

Research Article

Discrete Mittag-Leffler Functions in Linear Fractional Difference Equations

Jan Čermák, Tomáš Kisela, and Luděk Nechvátal

Institute of Mathematics, Faculty of Mechanical Engineering, Technická 2, 616 69 Brno, Czech Republic

Correspondence should be addressed to Jan Čermák, cermak.j@fme.vutbr.cz

Received 7 January 2011; Accepted 23 April 2011

Academic Editor: Yuri V. Rogovchenko

This paper investigates some initial value problems in discrete fractional calculus. We introduce a linear difference equation of fractional order along with suitable initial conditions of fractional type and prove the existence and uniqueness of the solution. Then the structure of the solutions space is discussed, and, in a particular case, an explicit form of the general solution involving discrete analogues of Mittag-Leffler functions is presented. All our observations are performed on a special time scale which unifies and generalizes ordinary difference calculus and q-difference calculus. Some of our results are new also in these particular discrete settings.

1. Introduction

The fractional calculus is a research field of mathematical analysis which may be taken for an old as well as a modern topic. It is an old topic because of its long history starting from some notes and ideas of G. W. Leibniz and L. Euler. On the other hand, it is a modern topic due to its enormous development during the last two decades. The present interest of many scientists and engineers in the theory of fractional calculus has been initiated by applications of this theory as well as by new mathematical challenges.

The theory of discrete fractional calculus belongs among these challenges. Foundations of this theory were formulated in pioneering works by Agarwal [1] and Diaz and Osler [2], where basic approaches, definitions, and properties of the theory of fractional sums and differences were reported (see also [3, 4]). The cited papers discussed these notions on discrete sets formed by arithmetic or geometric sequences (giving rise to fractional difference calculus or q-difference calculus). Recently, a series of papers continuing this research has appeared (see, e.g., [5, 6]).

The extension of basic notions of fractional calculus to other discrete settings was performed in [7], where fractional sums and differences have been introduced and studied in the framework of (q, h)-calculus, which can be reduced to ordinary difference calculus

to Contents, page 1

and q-difference calculus via the choice $q = h = 1$ and $h = 0$, respectively. This extension follows recent trends in continuous and discrete analysis, characterized by a unification and generalization, and resulting into the origin and progressive development of the time scales theory (see [8, 9]). Discussing problems of fractional calculus, a question concerning the introduction of (Hilger) fractional derivative or integral on arbitrary time scale turns out to be a difficult matter. Although first attempts have been already performed (see, e.g., [10]), results obtained in this direction seem to be unsatisfactory.

The aim of this paper is to introduce some linear nabla (q, h)-fractional difference equations (i.e., equations involving difference operators of noninteger orders) and investigate their basic properties. Some particular results concerning this topic are already known, either for ordinary difference equations or q-difference equations of fractional order (some relevant references will be mentioned in Section 4). We wish to unify them and also present results which are new even also in these particular discrete settings.

The structure of the paper is the following: Section 2 presents a necessary mathematical background related to discrete fractional calculus. In particular, we are going to make some general remarks concerning fractional calculus on arbitrary time scales. In Section 3, we consider a linear nabla (q, h)-difference equation of noninteger order and discuss the question of the existence and uniqueness of the solution for the corresponding initial value problem, as well as the question of a general solution of this equation. In Section 4, we consider a particular case of the studied equation and describe the base of its solutions space by the use of eigenfunctions of the corresponding difference operator. We show that these eigenfunctions can be taken for discrete analogues of the Mittag-Leffler functions.

2. Preliminaries

The basic definitions of fractional calculus on continuous or discrete settings usually originate from the Cauchy formula for repeated integration or summation, respectively. We state here its general form valid for arbitrary time scale $\mathbb{T}$. Before doing this, we recall the notion of Taylor monomials introduced in [9]. These monomials $\widehat{h}_n : \mathbb{T}^2 \to \mathbb{R}$, $n \in \mathbb{N}_0$ are defined recursively as follows:

$$\widehat{h}_0(t, s) = 1 \quad \forall s, t \in \mathbb{T} \tag{2.1}$$

and, given $\widehat{h}_n$ for $n \in \mathbb{N}_0$, we have

$$\widehat{h}_{n+1}(t, s) = \int_s^t \widehat{h}_n(\tau, s)\nabla\tau \quad \forall s, t \in \mathbb{T}\ . \tag{2.2}$$

Now let $f : \mathbb{T} \to \mathbb{R}$ be ∇-integrable on $[a, b] \cap \mathbb{T}$, $a,\ b \in \mathbb{T}$. We put

$${}_a\nabla^{-1} f(t) = \int_a^t f(\tau)\nabla\tau \quad \forall t \in \mathbb{T},\ a \le t \le b \tag{2.3}$$

to Contents, page 1

and define recursively

$$ {}_a\nabla^{-n} f(t) = \int_a^t {}_a\nabla^{-n+1} f(\tau)\nabla\tau \tag{2.4}$$

for $n = 2, 3, \ldots$. Then we have the following.

Proposition 2.1 (Nabla Cauchy formula). *Let $n \in \mathbb{Z}^+$, a, $b \in \mathbb{T}$ and let $f : \mathbb{T} \to \mathbb{R}$ be ∇-integrable on $[a, b] \cap \mathbb{T}$. If $t \in \mathbb{T}$, $a \le t \le b$, then*

$$ {}_a\nabla^{-n} f(t) = \int_a^t \widehat{h}_{n-1}\big(t, \rho(\tau)\big) f(\tau)\nabla\tau\ . \tag{2.5}$$

Proof. This assertion can be proved by induction. If $n = 1$, then (2.5) obviously holds. Let $n \ge 2$ and assume that (2.5) holds with n replaced with $n - 1$, that is,

$$ {}_a\nabla^{-n+1} f(t) = \int_a^t \widehat{h}_{n-2}\big(t, \rho(\tau)\big) f(\tau)\nabla\tau. \tag{2.6}$$

By the definition, the left-hand side of (2.5) is an antiderivative of ${}_a\nabla^{-n+1} f(t)$. We show that the right-hand side of (2.5) is an antiderivative of $\int_a^t \widehat{h}_{n-2}(t, \rho(\tau)) f(\tau)\nabla\tau$. Indeed, it holds

$$\nabla \int_a^t \widehat{h}_{n-1}\big(t, \rho(\tau)\big) f(\tau)\nabla\tau = \int_a^t \nabla\widehat{h}_{n-1}\big(t, \rho(\tau)\big) f(\tau)\nabla\tau = \int_a^t \widehat{h}_{n-2}\big(t, \rho(\tau)\big) f(\tau)\nabla\tau, \tag{2.7}$$

where we have employed the property

$$\nabla \int_a^t g(t, \tau)\nabla\tau = \int_a^t \nabla g(t, \tau)\nabla\tau + g\big(\rho(t), t\big) \tag{2.8}$$

(see [9, page 139]). Consequently, the relation (2.5) holds up to a possible additive constant. Substituting $t = a$, we can find this additive constant zero. □

The formula (2.5) is a corner stone in the introduction of the nabla fractional integral ${}_a\nabla^{-\alpha} f(t)$ for positive reals α. However, it requires a reasonable and natural extension of a discrete system of monomials $(\widehat{h}_n, n \in \mathbb{N}_0)$ to a continuous system $(\widehat{h}_\alpha, \alpha \in \mathbb{R}^+)$. This matter is closely related to a problem of an explicit form of $\widehat{h}_n$. Of course, it holds $\widehat{h}_1(t, s) = t - s$ for all $t, s \in \mathbb{T}$. However, the calculation of $\widehat{h}_n$ for $n > 1$ is a difficult task which seems to be answerable only in some particular cases. It is well known that for $\mathbb{T} = \mathbb{R}$, it holds

$$\widehat{h}_n(t, s) = \frac{(t - s)^n}{n!}, \tag{2.9}$$

to Contents, page 1

while for discrete time scales $\mathbb{T} = \mathbb{Z}$ and $\mathbb{T} = \overline{q^{\mathbb{Z}}} = \{q^k, k \in \mathbb{Z}\} \cup \{0\}$, $q > 1$, we have

$$\widehat{h}_n(t,s) = \frac{\prod_{j=0}^{n-1}(t-s+j)}{n!}, \qquad \widehat{h}_n(t,s) = \prod_{j=0}^{n-1}\frac{q^j t - s}{\sum_{r=0}^{j} q^r}, \tag{2.10}$$

respectively. In this connection, we recall a conventional notation used in ordinary difference calculus and q-calculus, namely,

$$(t-s)^{(n)} = \prod_{j=0}^{n-1}(t-s+j), \qquad (t-s)_{\widetilde{q}}^{(n)} = t^n \prod_{j=0}^{n-1}\left(1 - \frac{\widetilde{q}^j s}{t}\right) \quad (0 < \widetilde{q} < 1) \tag{2.11}$$

and $[j]_q = \sum_{r=0}^{j-1} q^r\, (q > 0)$, $[n]_q! = \prod_{j=1}^{n}[j]_q$. To extend the meaning of these symbols also for noninteger values (as it is required in the discrete fractional calculus), we recall some other necessary background of q-calculus. For any $x \in \mathbb{R}$ and $0 < q \neq 1$, we set $[x]_q = (q^x - 1)/(q-1)$. By the continuity, we put $[x]_1 = x$. Further, the q-Gamma function is defined for $0 < \widetilde{q} < 1$ as

$$\Gamma_{\widetilde{q}}(x) = \frac{(\widetilde{q},\widetilde{q})_\infty (1-\widetilde{q})^{1-x}}{(\widetilde{q}^x,\widetilde{q})_\infty}, \tag{2.12}$$

where $(p,\widetilde{q})_\infty = \prod_{j=0}^{\infty}(1 - p\widetilde{q}^j)$, $x \in \mathbb{R} \setminus \{0,-1,-2,\dots\}$. Note that this function satisfies the functional relation $\Gamma_{\widetilde{q}}(x+1) = [x]_{\widetilde{q}}\Gamma_{\widetilde{q}}(x)$ and the condition $\Gamma_{\widetilde{q}}(1) = 1$. Using this, the q-binomial coefficient can be introduced as

$$\begin{bmatrix} x \\ k \end{bmatrix}_{\widetilde{q}} = \frac{\Gamma_{\widetilde{q}}(x+1)}{\Gamma_{\widetilde{q}}(k+1)\Gamma_{\widetilde{q}}(x-k+1)}, \quad x \in \mathbb{R}, \quad k \in \mathbb{Z}. \tag{2.13}$$

Note that although the q-Gamma function is not defined at nonpositive integers, the formula

$$\frac{\Gamma_{\widetilde{q}}(x+m)}{\Gamma_{\widetilde{q}}(x)} = (-1)^m \widetilde{q}^{\,xm+\binom{m}{2}} \frac{\Gamma_{\widetilde{q}}(1-x)}{\Gamma_{\widetilde{q}}(1-x-m)}, \quad x \in \mathbb{R},\ m \in \mathbb{Z}^+ \tag{2.14}$$

permits to calculate this ratio also at such the points. It is well known that if $\widetilde{q} \to 1^-$ then $\Gamma_{\widetilde{q}}(x)$ becomes the Euler Gamma function $\Gamma(x)$ (and analogously for the q-binomial coefficient). Among many interesting properties of the q-Gamma function and q-binomial coefficients, we mention q-Pascal rules

$$\begin{bmatrix} x \\ k \end{bmatrix}_{\widetilde{q}} = \begin{bmatrix} x-1 \\ k-1 \end{bmatrix}_{\widetilde{q}} + \widetilde{q}^k \begin{bmatrix} x-1 \\ k \end{bmatrix}_{\widetilde{q}}, \quad x \in \mathbb{R},\ k \in \mathbb{Z}\,, \tag{2.15}$$

$$\begin{bmatrix} x \\ k \end{bmatrix}_{\widetilde{q}} = \widetilde{q}^{\,x-k} \begin{bmatrix} x-1 \\ k-1 \end{bmatrix}_{\widetilde{q}} + \begin{bmatrix} x-1 \\ k \end{bmatrix}_{\widetilde{q}}, \quad x \in \mathbb{R},\ k \in \mathbb{Z} \tag{2.16}$$

to Contents, page 1

and the q-Vandermonde identity

$$\sum_{j=0}^{m} \begin{bmatrix} x \\ m-j \end{bmatrix}_{\tilde{q}} \begin{bmatrix} y \\ j \end{bmatrix}_{\tilde{q}} \tilde{q}^{j^2-mj+xj} = \begin{bmatrix} x+y \\ m \end{bmatrix}_{\tilde{q}}, \quad x, y \in \mathbb{R}, \ m \in \mathbb{N}_0 \tag{2.17}$$

(see [11]) that turn out to be very useful in our further investigations.

The computation of an explicit form of $\widehat{h}_n(t,s)$ can be performed also in a more general case. We consider here the time scale

$$\mathbb{T}^{t_0}_{(q,h)} = \left\{ t_0 q^k + [k]_q h, k \in \mathbb{Z} \right\} \cup \left\{ \frac{h}{1-q} \right\}, \quad t_0 > 0, \ q \geq 1, \ h \geq 0, \ q+h > 1 \tag{2.18}$$

(see also [7]). Note that if $q = 1$ then the cluster point $h/(1-q) = -\infty$ is not involved in $\mathbb{T}^{t_0}_{(q,h)}$. The forward and backward jump operator is the linear function $\sigma(t) = qt + h$ and $\rho(t) = q^{-1}(t-h)$, respectively. Similarly, the forward and backward graininess is given by $\mu(t) = (q-1)t + h$ and $\nu(t) = q^{-1}\mu(t)$, respectively. In particular, if $t_0 = q = h = 1$, then $\mathbb{T}^{t_0}_{(q,h)}$ becomes $\mathbb{Z}$, and if $t_0 = 1$, $q > 1$, $h = 0$, then $\mathbb{T}^{t_0}_{(q,h)}$ is reduced to $\overline{q^{\mathbb{Z}}}$.

Let $a \in \mathbb{T}^{t_0}_{(q,h)}$, $a > h/(1-q)$ be fixed. Then we introduce restrictions of the time scale $\mathbb{T}^{t_0}_{(q,h)}$ by the relation

$$\widetilde{\mathbb{T}}^{\sigma^i(a)}_{(q,h)} = \left\{ t \in \mathbb{T}^{t_0}_{(q,h)}, t \geq \sigma^i(a) \right\}, \quad i = 0, 1, \ldots, \tag{2.19}$$

where the symbol σ^i stands for the ith iterate of σ (analogously, we use the symbol ρ^i). To simplify the notation, we put $\tilde{q} = 1/q$ whenever considering the time scale $\mathbb{T}^{t_0}_{(q,h)}$ or $\widetilde{\mathbb{T}}^{\sigma^i(a)}_{(q,h)}$.

Using the induction principle, we can verify that Taylor monomials on $\mathbb{T}^{t_0}_{(q,h)}$ have the form

$$\widehat{h}_n(t,s) = \frac{\prod_{j=0}^{n-1} \left(\sigma^j(t) - s \right)}{[n]_q!} = \frac{\prod_{j=0}^{n-1} \left(t - \rho^j(s) \right)}{[n]_{\tilde{q}}!}. \tag{2.20}$$

Note that this result generalizes previous forms (2.10) and, moreover, enables its unified notation. In particular, if we introduce the symbolic (q,h)-power

$$(t-s)^{(n)}_{(\tilde{q},h)} = \prod_{j=0}^{n-1} \left(t - \rho^j(s) \right) \tag{2.21}$$

unifying (2.11), then the Cauchy formula (2.5) can be rewritten for $\mathbb{T} = \mathbb{T}^{t_0}_{(q,h)}$ as

$${}_a\nabla^{-n} f(t) = \int_a^t \frac{\left(t - \rho(\tau) \right)^{(n-1)}_{(\tilde{q},h)}}{[n-1]_{\tilde{q}}!} f(\tau) \nabla\tau. \tag{2.22}$$

to Contents, page 1

Discussing a reasonable generalization of (q,h)-power (2.21) to real values α instead of integers n, we recall broadly accepted extensions of its particular cases (2.11) in the form

$$(t-s)^{(\alpha)}=\frac{\Gamma(t-s+\alpha)}{\Gamma(t-s)},\qquad (t-s)^{(\alpha)}_{\tilde{q}}=t^{\alpha}\frac{(s/t,\tilde{q})_{\infty}}{(\tilde{q}^{\alpha}s/t,\tilde{q})_{\infty}},\qquad t\neq 0. \tag{2.23}$$

Now, we assume $s,t\in\mathbb{T}^{t_0}_{(q,h)}$, $t\geq s>h/(1-q)$. First, consider (q,h)-power (2.21) corresponding to the time scale $\mathbb{T}^{t_0}_{(q,h)}$, where $q>1$. Then we can rewrite (2.21) as

$$(t-s)^{(n)}_{(\tilde{q},h)}=\left(t+\frac{h\tilde{q}}{1-\tilde{q}}\right)^{n}\prod_{j=0}^{n-1}\left(1-\tilde{q}^{j}\frac{s+h\tilde{q}/(1-\tilde{q})}{t+h\tilde{q}/(1-\tilde{q})}\right)=([t]-[s])^{(n)}_{\tilde{q}}, \tag{2.24}$$

where $[t]=t+h\tilde{q}/(1-\tilde{q})$ and $[s]=s+h\tilde{q}/(1-\tilde{q})$. A required extension of (q,h)-power (2.21) is then provided by the formula

$$(t-s)^{(\alpha)}_{(\tilde{q},h)}=([t]-[s])^{(\alpha)}_{\tilde{q}}. \tag{2.25}$$

Now consider (q,h)-power (2.21) corresponding to the time scale $\mathbb{T}^{t_0}_{(q,h)}$, where $q=1$. Then

$$(t-s)^{(n)}_{(1,h)}=\prod_{j=0}^{n-1}(t-s+jh)=h^{n}\prod_{j=0}^{n-1}\left(\frac{t-s}{h}+j\right)=h^{n}\frac{((t-s)/h+n-1)!}{((t-s)/h-1)!} \tag{2.26}$$

and the formula (2.21) can be extended by

$$(t-s)^{(\alpha)}_{(1,h)}=\frac{h^{\alpha}\Gamma((t-s)/h+\alpha)}{\Gamma((t-s)/h)}. \tag{2.27}$$

These definitions are consistent, since it can be shown that

$$\lim_{\tilde{q}\to 1^{-}}([t]-[s])^{(\alpha)}_{\tilde{q}}=(t-s)^{(\alpha)}_{(1,h)}. \tag{2.28}$$

Now the required extension of the monomial $\widehat{h}_n(t,s)$ corresponding to $\mathbb{T}^{t_0}_{(q,h)}$ takes the form

$$\widehat{h}_{\alpha}(t,s)=\frac{(t-s)^{(\alpha)}_{(\tilde{q},h)}}{\Gamma_{\tilde{q}}(\alpha+1)}. \tag{2.29}$$

Another (equivalent) expression of $\widehat{h}_{\alpha}(t,s)$ is provided by the following assertion.

to Contents, page 1

Proposition 2.2. *Let $\alpha \in \mathbb{R}$, $s, t \in \mathbb{T}_{(q,h)}^{t_0}$ and $n \in \mathbb{N}_0$ be such that $t = \sigma^n(s)$. Then*

$$\widehat{h}_\alpha(t,s) = (\nu(t))^\alpha \begin{bmatrix} \alpha + n - 1 \\ n - 1 \end{bmatrix}_{\widetilde{q}} = (\nu(t))^\alpha \begin{bmatrix} -\alpha - 1 \\ n - 1 \end{bmatrix}_{\widetilde{q}} (-1)^{n-1} \widetilde{q}^{\alpha(n-1)+\binom{n}{2}}. \tag{2.30}$$

Proof. Let $q > 1$. Using the relations

$$[t] = \frac{\nu(t)}{(1-\widetilde{q})}, \qquad \frac{[s]}{[t]} = \widetilde{q}^n, \tag{2.31}$$

we can derive that

$$\begin{aligned} \widehat{h}_\alpha(t,s) &= \frac{[t]^\alpha([s]/[t],\widetilde{q})_\infty}{\Gamma_{\widetilde{q}}(\alpha+1)(\widetilde{q}^\alpha[s]/[t],\widetilde{q})_\infty} = \frac{(1-\widetilde{q})^{-\alpha}\nu(t)^\alpha(\widetilde{q}^n,\widetilde{q})_\infty}{\Gamma_{\widetilde{q}}(\alpha+1)(\widetilde{q}^{\alpha+n},\widetilde{q})_\infty} \\ &= (\nu(t))^\alpha \frac{\Gamma_{\widetilde{q}}(\alpha+n)}{\Gamma_{\widetilde{q}}(\alpha+1)\Gamma_{\widetilde{q}}(n)} = (\nu(t))^\alpha \begin{bmatrix} \alpha + n - 1 \\ n - 1 \end{bmatrix}_{\widetilde{q}}. \end{aligned} \tag{2.32}$$

The second equality in (2.30) follows from the identity (2.14). The case $q = 1$ results from (2.27). □

The key property of $\widehat{h}_\alpha(t,s)$ follows from its differentiation. The symbol $\nabla_{(q,h)}^m$ used in the following assertion (and also undermentioned) is the mth order nabla (q,h)-derivative on the time scale $\mathbb{T}_{(q,h)}^{t_0}$, defined for $m = 1$ as

$$\nabla_{(q,h)} f(t) = \frac{f(t) - f(\rho(t))}{\nu(t)} = \frac{f(t) - f(\widetilde{q}(t-h))}{(1-\widetilde{q})t + \widetilde{q}h} \tag{2.33}$$

and iteratively for higher orders.

Lemma 2.3. *Let $m \in \mathbb{Z}^+$, $\alpha \in \mathbb{R}$, $s, t \in \mathbb{T}_{(q,h)}^{t_0}$ and $n \in \mathbb{Z}^+$, $n \geq m$ be such that $t = \sigma^n(s)$. Then*

$$\nabla_{(q,h)}^m \widehat{h}_\alpha(t,s) = \begin{cases} \widehat{h}_{\alpha-m}(t,s), & \alpha \notin \{0,1,\ldots,m-1\}, \\ 0, & \alpha \in \{0,1,\ldots,m-1\}. \end{cases} \tag{2.34}$$

to Contents, page 1

Proof. First let $m = 1$. For $\alpha = 0$ we get $\widehat{h}_0(t,s) = 1$ and the first nabla (q,h)-derivative is zero. If $\alpha \neq 0$, then by (2.30) and (2.16), we have

$$\begin{aligned}
\nabla_{(q,h)}\widehat{h}_\alpha(t,s) &= \frac{\widehat{h}_\alpha(t,s) - \widehat{h}_\alpha(\rho(t),s)}{\nu(t)} \\
&= \frac{1}{\nu(t)}\left((\nu(t))^\alpha \begin{bmatrix} \alpha+n-1 \\ n-1 \end{bmatrix}_{\tilde{q}} - (\nu(\rho(t)))^\alpha \begin{bmatrix} \alpha+n-2 \\ n-2 \end{bmatrix}_{\tilde{q}} \right) \qquad (2.35) \\
&= (\nu(t))^{\alpha-1}\left(\begin{bmatrix} \alpha+n-1 \\ n-1 \end{bmatrix}_{\tilde{q}} - \tilde{q}^\alpha \begin{bmatrix} \alpha+n-2 \\ n-2 \end{bmatrix}_{\tilde{q}} \right) = \widehat{h}_{\alpha-1}(t,s).
\end{aligned}$$

The case $m \geq 2$ can be verified by the induction principle. □

We note that an extension of this property for derivatives of noninteger orders will be performed in Section 4.

Now we can continue with the introduction of (q,h)-fractional integral and derivative of a function $f : \widetilde{\mathbb{T}}^a_{(q,h)} \to \mathbb{R}$. Let $t \in \widetilde{\mathbb{T}}^a_{(q,h)}$. Our previous considerations (in particular, the Cauchy formula (2.5) along with the relations (2.22) and (2.29)) warrant us to introduce the nabla (q,h)-fractional integral of order $\alpha \in \mathbb{R}^+$ over the time scale interval $[a,t] \cap \widetilde{\mathbb{T}}^a_{(q,h)}$ as

$$ {}_a\nabla^{-\alpha}_{(q,h)} f(t) = \int_a^t \widehat{h}_{\alpha-1}(t,\rho(\tau)) f(\tau)\nabla\tau \qquad (2.36)$$

(see also [7]). The nabla (q,h)-fractional derivative of order $\alpha \in \mathbb{R}^+$ is then defined by

$$ {}_a\nabla^{\alpha}_{(q,h)} f(t) = \nabla^m_{(q,h)}\, {}_a\nabla^{-(m-\alpha)}_{(q,h)} f(t), \qquad (2.37)$$

where $m \in \mathbb{Z}^+$ is given by $m-1 < \alpha \leq m$. For the sake of completeness, we put

$$ {}_a\nabla^{0}_{(q,h)} f(t) = f(t). \qquad (2.38)$$

As we noted earlier, a reasonable introduction of fractional integrals and fractional derivatives on arbitrary time scales remains an open problem. In the previous part, we have consistently used (and in the sequel, we shall consistently use) the time scale notation of main procedures and operations to outline a possible way out to further generalizations.

3. A Linear Initial Value Problem

In this section, we are going to discuss the linear initial value problem

$$\sum_{j=1}^m p_{m-j+1}(t)\, {}_a\nabla^{\alpha-j+1}_{(q,h)} y(t) + p_0(t)y(t) = 0\,, \qquad t \in \widetilde{\mathbb{T}}^{\sigma^{m+1}(a)}_{(q,h)}, \qquad (3.1)$$

to Contents, page 1

$$ {}_a\nabla^{\alpha-j}_{(q,h)} y(t)\Big|_{t=\sigma^m(a)} = y_{\alpha-j}, \quad j = 1,2,\ldots,m, \tag{3.2}$$

where $\alpha \in \mathbb{R}^+$ and $m \in \mathbb{Z}^+$ are such that $m-1 < \alpha \leq m$. Further, we assume that $p_j(t)$ are arbitrary real-valued functions on $\widetilde{\mathbb{T}}^{\sigma^{m+1}(a)}_{(q,h)}$ $(j = 1,\ldots,m-1)$, $p_m(t) = 1$ on $\widetilde{\mathbb{T}}^{\sigma^{m+1}(a)}_{(q,h)}$ and $y_{\alpha-j}$ $(j = 1,\ldots,m)$ are arbitrary real scalars.

If α is a positive integer, then (3.1)-(3.2) becomes the standard discrete initial value problem. If α is not an integer, then applying the definition of nabla (q,h)-fractional derivatives, we can observe that (3.1) is of the general form

$$\sum_{i=0}^{n-1} a_i(t) y\left(\rho^i(t)\right) = 0\ , \quad t \in \widetilde{\mathbb{T}}^{\sigma^{m+1}(a)}_{(q,h)},\ n \text{ being such that } t = \sigma^n(a), \tag{3.3}$$

which is usually referred to as the equation of Volterra type. If such an equation has two different solutions, then their values differ at least at one of the points $\sigma(a), \sigma^2(a), \ldots, \sigma^m(a)$. In particular, if $a_0(t) \neq 0$ for all $t \in \widetilde{\mathbb{T}}^{\sigma^{m+1}(a)}_{(q,h)}$, then arbitrary values of $y(\sigma(a)), y(\sigma^2(a)), \ldots, y(\sigma^m(a))$ determine uniquely the solution $y(t)$ for all $t \in \widetilde{\mathbb{T}}^{\sigma^{m+1}(a)}_{(q,h)}$. We show that the values $y_{\alpha-1}, y_{\alpha-2}, \ldots, y_{\alpha-m}$, introduced by (3.2), keep the same properties.

Proposition 3.1. *Let $y : \widetilde{\mathbb{T}}^{\sigma(a)}_{(q,h)} \to \mathbb{R}$ be a function. Then* (3.2) *represents a one-to-one mapping between the vectors $(y(\sigma(a)), y(\sigma^2(a)), \ldots, y(\sigma^m(a)))$ and $(y_{\alpha-1}, y_{\alpha-2}, \ldots, y_{\alpha-m})$.*

Proof. The case $\alpha \in \mathbb{Z}^+$ is well known from the literature. Let $\alpha \notin \mathbb{Z}^+$. We wish to show that the values of $y(\sigma(a)), y(\sigma^2(a)), \ldots, y(\sigma^m(a))$ determine uniquely the values of

$$ {}_a\nabla^{\alpha-1}_{(q,h)} y(t)\Big|_{t=\sigma^m(a)}, \quad {}_a\nabla^{\alpha-2}_{(q,h)} y(t)\Big|_{t=\sigma^m(a)}, \ldots, {}_a\nabla^{\alpha-m}_{(q,h)} y(t)\Big|_{t=\sigma^m(a)} \tag{3.4}$$

and vice versa. Utilizing the relation

$$ {}_a\nabla^{\alpha-j}_{(q,h)} y(t)\Big|_{t=\sigma^m(a)} = \sum_{k=1}^{m} \nu\left(\sigma^{m-k+1}(a)\right) \widehat{h}_{j-1-\alpha}\left(\sigma^m(a), \sigma^{m-k}(a)\right) y\left(\sigma^{m-k+1}(a)\right) \tag{3.5}$$

(see [7, Propositions 1 and 3] with respect to (2.30)), we can rewrite (3.2) as the linear mapping

$$\sum_{k=1}^{m} r_{jk} y\left(\sigma^{m-k+1}(a)\right) = y_{\alpha-j}, \quad j = 1,\ldots,m, \tag{3.6}$$

where

$$r_{jk} = \nu\left(\sigma^{m-k+1}(a)\right) \widehat{h}_{j-1-\alpha}\left(\sigma^m(a), \sigma^{m-k}(a)\right), \quad j,k = 1,\ldots,m \tag{3.7}$$

to Contents, page 1

are elements of the transformation matrix R_m. We show that R_m is regular. Obviously,

$$\det R_m = \left(\prod_{k=1}^{m} \nu\left(\sigma^k(a)\right)\right) \det H_m, \tag{3.8}$$

where

$$H_m = \begin{pmatrix} \widehat{h}_{-\alpha}(\sigma^m(a),\sigma^{m-1}(a)) & \widehat{h}_{-\alpha}(\sigma^m(a),\sigma^{m-2}(a)) & \cdots & \widehat{h}_{-\alpha}(\sigma^m(a),a) \\ \widehat{h}_{1-\alpha}(\sigma^m(a),\sigma^{m-1}(a)) & \widehat{h}_{1-\alpha}(\sigma^m(a),\sigma^{m-2}(a)) & \cdots & \widehat{h}_{1-\alpha}(\sigma^m(a),a) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{h}_{m-1-\alpha}(\sigma^m(a),\sigma^{m-1}(a)) & \widehat{h}_{m-1-\alpha}(\sigma^m(a),\sigma^{m-2}(a)) & \cdots & \widehat{h}_{m-1-\alpha}(\sigma^m(a),a) \end{pmatrix}. \tag{3.9}$$

To calculate $\det H_m$, we employ some elementary operations preserving the value of $\det H_m$. Using the properties

$$\begin{aligned} &\widehat{h}_{i-\alpha}\left(\sigma^m(a),\sigma^\ell(a)\right) - \nu(\sigma^m(a))\widehat{h}_{i-\alpha-1}\left(\sigma^m(a),\sigma^\ell(a)\right) = \widehat{h}_{i-\alpha}\left(\sigma^{m-1}(a),\sigma^\ell(a)\right) \\ &\qquad\qquad (i = 1,2,\ldots,m-1,\ l = 0,1,\ldots m-2), \qquad (3.10) \\ &\widehat{h}_{i-\alpha}\left(\sigma^m(a),\sigma^{m-1}(a)\right) - \nu(\sigma^m(a))\widehat{h}_{i-\alpha-1}\left(\sigma^m(a),\sigma^{m-1}(a)\right) = 0, \end{aligned}$$

which follow from Lemma 2.3, we multiply the ith row $(i = 1,2,\ldots,m-1)$ of H_m by $-\nu(\sigma^m(a))$ and add it to the successive one. We arrive at the form

$$\left(\begin{array}{c|ccc} \widehat{h}_{-\alpha}(\sigma^m(a),\sigma^{m-1}(a)) & \widehat{h}_{-\alpha}(\sigma^m(a),\sigma^{m-2}(a)) & \cdots & \widehat{h}_{-\alpha}(\sigma^m(a),a) \\ \hline 0 & & & \\ \vdots & & H_{m-1} & \\ 0 & & & \end{array}\right). \tag{3.11}$$

Then we apply repeatedly this procedure to obtain the triangular matrix

$$\begin{pmatrix} \widehat{h}_{-\alpha}(\sigma^m(a),\sigma^{m-1}(a)) & \widehat{h}_{-\alpha}(\sigma^m(a),\sigma^{m-2}(a)) & \cdots & \widehat{h}_{-\alpha}(\sigma^m(a),a) \\ 0 & \widehat{h}_{1-\alpha}(\sigma^{m-1}(a),\sigma^{m-2}(a)) & \cdots & \widehat{h}_{1-\alpha}(\sigma^{m-1}(a),a) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \widehat{h}_{m-1-\alpha}(\sigma(a),a) \end{pmatrix}. \tag{3.12}$$

Since $\widehat{h}_{i-\alpha}(\sigma^k(a),\sigma^{k-1}(a)) = (\nu(\sigma^k(a))^{i-\alpha}(i = 0,1,\ldots,m-1)$, we get

to Contents, page 1

$$\det H_m = \prod_{k=1}^{m}\left(\nu\left(\sigma^k(a)\right)\right)^{m-k-\alpha}, \text{ that is,} \quad \det R_m = \prod_{k=1}^{m}\left(\nu\left(\sigma^k(a)\right)\right)^{m-k-\alpha+1} \neq 0. \tag{3.13}$$

Thus the matrix R_m is regular, hence the corresponding mapping (3.6) is one to one. □

Now we approach a problem of the existence and uniqueness of (3.1)-(3.2). First we recall the general notion of ν-regressivity of a matrix function and a corresponding linear nabla dynamic system (see [9]).

Definition 3.2. An $n \times n$-matrix-valued function $A(t)$ on a time scale $\mathbb{T}$ is called ν-regressive provided

$$\det(I - \nu(t)A(t)) \neq 0 \quad \forall t \in \mathbb{T}_\kappa, \tag{3.14}$$

where I is the identity matrix. Further, we say that the linear dynamic system

$$\nabla z(t) = A(t)z(t) \tag{3.15}$$

is ν-regressive provided that $A(t)$ is ν-regressive.

Considering a higher order linear difference equation, the notion of ν-regressivity for such an equation can be introduced by means of its transformation to the corresponding first order linear dynamic system. We are going to follow this approach and generalize the notion of ν-regressivity for the linear fractional difference equation (3.1).

Definition 3.3. Let $\alpha \in \mathbb{R}^+$ and $m \in \mathbb{Z}^+$ be such that $m-1 < \alpha \leq m$. Then (3.1) is called ν-regressive provided the matrix

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -\dfrac{p_0(t)}{\nu^{m-\alpha}(t)} & -p_1(t) & \cdots & -p_{m-2}(t) & -p_{m-1}(t) \end{pmatrix} \tag{3.16}$$

is ν-regressive.

Remark 3.4. The explicit expression of the ν-regressivity property for (3.1) can be read as

$$1 + \sum_{j=1}^{m-1} p_{m-j}(t)(\nu(t))^j + p_0(t)(\nu(t))^\alpha \neq 0 \quad \forall t \in \widetilde{\mathbb{T}}_{(q,h)}^{\sigma^{m+1}(a)}. \tag{3.17}$$

If α is a positive integer, then both these introductions agree with the definition of ν-regressivity of a higher order linear difference equation presented in [9].

to Contents, page 1

Theorem 3.5. *Let* (3.1) *be* ν*-regressive. Then the problem* (3.1)-(3.2) *has a unique solution defined for all* $t \in \widetilde{\mathbb{T}}_{(q,h)}^{\sigma(a)}$.

Proof. The conditions (3.2) enable us to determine the values of $y(\sigma(a)), y(\sigma^2(a)), \ldots, y(\sigma^m(a))$ by the use of (3.6). To calculate the values of $y(\sigma^{m+1}(a)), y(\sigma^{m+2}(a)), \ldots,$ we perform the transformation

$$z_j(t) = {}_a\nabla_{(q,h)}^{\alpha-m+j-1} y(t), \quad t \in \widetilde{\mathbb{T}}_{(q,h)}^{\sigma^j(a)}, \; j = 1, 2, \ldots, m \tag{3.18}$$

which allows us to rewrite (3.1) into a matrix form. Before doing this, we need to express $y(t)$ in terms of $z_1(t),\; z_1(\rho(t)), \ldots, z_1(\sigma(a))$. Applying the relation ${}_a\nabla_{(q,h)}^{m-\alpha}\, {}_a\nabla_{(q,h)}^{-(m-\alpha)} y(t) = y(t)$ (see [7]) and expanding the fractional derivative, we arrive at

$$y(t) = {}_a\nabla_{(q,h)}^{m-\alpha} z_1(t) = \frac{z_1(t)}{\nu^{m-\alpha}(t)} + \int_a^{\rho(t)} \widehat{h}_{\alpha-m-1}(t, \rho(\tau)) z_1(\tau) \nabla\tau. \tag{3.19}$$

Therefore, the problem (3.1)-(3.2) can be rewritten to the vector form

$$\begin{gathered} {}_a\nabla_{(q,h)} z(t) = A(t) z(t) + b(t), \quad t \in \widetilde{\mathbb{T}}_{(q,h)}^{\sigma^{m+1}(a)}, \\ z(\sigma^m(a)) = \left(y_{\alpha-m}, \ldots, y_{\alpha-1}\right)^T, \end{gathered} \tag{3.20}$$

where

$$z(t) = (z_1(t), \ldots, z_m(t))^T, \qquad b(t) = \left(0, \ldots, 0, -p_0(t) \int_a^{\rho(t)} \widehat{h}_{\alpha-m-1}(t, \rho(\tau)) z_1(\tau) \nabla\tau \right)^T \tag{3.21}$$

and $A(t)$ is given by (3.16). The ν-regressivity of the matrix $A(t)$ enables us to write

$$z(t) = (I - \nu(t) A(t))^{-1} \left(z(\rho(t)) + \nu(t) b(t) \right), \quad t \in \widetilde{\mathbb{T}}_{(q,h)}^{\sigma^{m+1}(a)}, \tag{3.22}$$

hence, using the value of $z(\sigma^m(a))$, we can solve this system by the step method starting from $t = \sigma^{m+1}(a)$. The solution $y(t)$ of the original initial value problem (3.1)-(3.2) is then given by the formula (3.19). □

Remark 3.6. The previous assertion on the existence and uniqueness of the solution can be easily extended to the initial value problem involving nonhomogeneous linear equations as well as some nonlinear equations.

The final goal of this section is to investigate the structure of the solutions of (3.1). We start with the following notion.

Definition 3.7. Let $\gamma \in \mathbb{R}$, $0 \leq \gamma < 1$. For m functions $y_j : \widetilde{\mathbb{T}}^{a}_{(q,h)} \to \mathbb{R}(j = 1, 2, \ldots, m)$, we define the γ-Wronskian $W_\gamma(y_1, \ldots, y_m)(t)$ as determinant of the matrix

$$V_\gamma(y_1, \ldots, y_m)(t) = \begin{pmatrix} {}_a\nabla^{-\gamma}_{(q,h)} y_1(t) & {}_a\nabla^{-\gamma}_{(q,h)} y_2(t) & \cdots & {}_a\nabla^{-\gamma}_{(q,h)} y_m(t) \\ {}_a\nabla^{1-\gamma}_{(q,h)} y_1(t) & {}_a\nabla^{1-\gamma}_{(q,h)} y_2(t) & \cdots & {}_a\nabla^{1-\gamma}_{(q,h)} y_m(t) \\ \vdots & \vdots & \ddots & \vdots \\ {}_a\nabla^{m-1-\gamma}_{(q,h)} y_1(t) & {}_a\nabla^{m-1-\gamma}_{(q,h)} y_2(t) & \cdots & {}_a\nabla^{m-1-\gamma}_{(q,h)} y_m(t) \end{pmatrix}, \quad t \in \widetilde{\mathbb{T}}^{\sigma^m(a)}_{(q,h)}. \tag{3.23}$$

Remark 3.8. Note that the first row of this matrix involves fractional order integrals. It is a consequence of the form of initial conditions utilized in our investigations. Of course, this introduction of $W_\gamma(y_1, \ldots, y_m)(t)$ coincides for $\gamma = 0$ with the classical definition of the Wronskian (see [8]). Moreover, it holds $W_\gamma(y_1, \ldots, y_m)(t) = W_0({}_a\nabla^{-\gamma}_{(q,h)} y_1, \ldots, {}_a\nabla^{-\gamma}_{(q,h)} y_m)(t)$.

Theorem 3.9. *Let functions $y_1(t), \ldots, y_m(t)$ be solutions of the ν-regressive equation* (3.1) *and let $W_{m-\alpha}(y_1, \ldots, y_m)(\sigma^m(a)) \neq 0$. Then any solution $y(t)$ of* (3.1) *can be written in the form*

$$y(t) = \sum_{k=1}^{m} c_k y_k(t), \quad t \in \widetilde{\mathbb{T}}^{\sigma(a)}_{(q,h)}, \tag{3.24}$$

where $c_1, \ldots, c_m$ are real constants.

Proof. Let $y(t)$ be a solution of (3.1). By Proposition 3.1, there exist real scalars $y_{\alpha-1}, \ldots, y_{\alpha-m}$ such that $y(t)$ is satisfying (3.2). Now we consider the function $u(t) = \sum_{k=1}^{m} c_k y_k(t)$, where the m-tuple $(c_1, \ldots, c_m)$ is the unique solution of

$$V_{m-\alpha}(y_1, \ldots, y_m)(\sigma^m(a)) \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} y_{\alpha-m} \\ y_{\alpha-m+1} \\ \vdots \\ y_{\alpha-1} \end{pmatrix}. \tag{3.25}$$

The linearity of (3.1) implies that $u(t)$ has to be its solution. Moreover, it holds

$${}_a\nabla^{\alpha-j}_{(q,h)} u(t)\Big|_{t=\sigma^m(a)} = y_{\alpha-j}, \quad j = 1, 2, \ldots, m, \tag{3.26}$$

hence $u(t)$ is a solution of the initial value problem (3.1)-(3.2). By Theorem 3.5, it must be $y(t) = u(t)$ for all $t \in \widetilde{\mathbb{T}}^{\sigma(a)}_{(q,h)}$ and (3.24) holds. $\square$

Remark 3.10. The formula (3.24) is essentially an expression of the general solution of (3.1).

to Contents, page 1

4. Two-Term Equation and (q, h)-Mittag-Leffler Function

Our main interest in this section is to find eigenfunctions of the fractional operator ${}_a\nabla^{\alpha}_{(q,h)}$, $\alpha \in \mathbb{R}^+$. In other words, we wish to solve (3.1) in a special form

$$ {}_a\nabla^{\alpha}_{(q,h)} y(t) = \lambda y(t), \quad \lambda \in \mathbb{R}, \quad t \in \widetilde{\mathbb{T}}^{\sigma^{m+1}(a)}_{(q,h)}. \tag{4.1} $$

Throughout this section, we assume that ν-regressivity condition for (4.1) is ensured, that is,

$$ \lambda(\nu(t))^{\alpha} \neq 1. \tag{4.2} $$

Discussions on methods of solving fractional difference equations are just at the beginning. Some techniques how to explicitly solve these equations (at least in particular cases) are exhibited, for example, in [12–14], where a discrete analogue of the Laplace transform turns out to be the most developed method. In this section, we describe the technique not utilizing the transform method, but directly originating from the role which is played by the Mittag-Leffler function in the continuous fractional calculus (see, e.g., [15]). In particular, we introduce the notion of a discrete Mittag-Leffler function in a setting formed by the time scale $\widetilde{\mathbb{T}}^{a}_{(q,h)}$ and demonstrate its significance with respect to eigenfunctions of the operator ${}_a\nabla^{\alpha}_{(q,h)}$. These results generalize and extend those derived in [16, 17].

We start with the power rule stated in Lemma 2.3 and perform its extension to fractional integrals and derivatives.

Proposition 4.1. *Let $\alpha \in \mathbb{R}^+$, $\beta \in \mathbb{R}$ and $t \in \widetilde{\mathbb{T}}^{\sigma(a)}_{(q,h)}$. Then it holds*

$$ {}_a\nabla^{-\alpha}_{(q,h)} \widehat{h}_{\beta}(t, a) = \widehat{h}_{\alpha+\beta}(t, a). \tag{4.3} $$

Proof. Let $t \in \widetilde{\mathbb{T}}^{\sigma(a)}_{(q,h)}$ be such that $t = \sigma^n(a)$ for some $n \in \mathbb{Z}^+$. We have

$$ \begin{aligned} {}_a\nabla^{-\alpha}_{(q,h)} \widehat{h}_{\beta}(t, a) &= \sum_{k=0}^{n-1} \widehat{h}_{\alpha-1}\left(t, \rho^{k+1}(t)\right) \nu\left(\rho^k(t)\right) \widehat{h}_{\beta}\left(\rho^k(t), a\right) \\ &= \sum_{k=0}^{n-1} (\nu(t))^{\alpha-1} \begin{bmatrix} -\alpha \\ k \end{bmatrix}_{\widetilde{q}} (-1)^k \widetilde{q}^{(\alpha-1)k+\binom{k+1}{2}} \widetilde{q}^k \nu(t) \\ &\quad \times \left(\nu\left(\rho^k(t)\right)\right)^{\beta} \begin{bmatrix} -\beta-1 \\ n-k-1 \end{bmatrix}_{\widetilde{q}} (-1)^{n-k-1} \widetilde{q}^{\beta(n-k-1)+\binom{n-k}{2}} \\ &= (\nu(t))^{\alpha+\beta} \sum_{k=0}^{n-1} \begin{bmatrix} -\alpha \\ k \end{bmatrix}_{\widetilde{q}} \begin{bmatrix} -\beta-1 \\ n-k-1 \end{bmatrix}_{\widetilde{q}} (-1)^{n-1} \widetilde{q}^{k^2-k(n-1)+k\alpha+\binom{n}{2}+\beta(n-1)} \end{aligned} $$

to Contents, page 1

$$
\begin{aligned}
&= (\nu(t))^{\alpha+\beta}\sum_{k=0}^{n-1}\begin{bmatrix} -\alpha \\ n-k-1\end{bmatrix}_{\tilde{q}}\begin{bmatrix} -\beta-1 \\ k\end{bmatrix}_{\tilde{q}} \\
&\qquad\times (-1)^{n-1}\tilde{q}^{(n-k-1)^2-(n-k-1)(n-1)+(n-k-1)\alpha+\binom{n}{2}+\beta(n-1)} \\
&= (\nu(t))^{\alpha+\beta}\sum_{k=0}^{n-1}\begin{bmatrix} -\alpha \\ n-k-1\end{bmatrix}_{\tilde{q}}\begin{bmatrix} -\beta-1 \\ k\end{bmatrix}_{\tilde{q}}(-1)^{n-1}\tilde{q}^{k^2-k(n-1)-k\alpha+(\alpha+\beta)(n-1)+\binom{n}{2}} \\
&= (\nu(t))^{\alpha+\beta}\begin{bmatrix} -\alpha-\beta-1 \\ n-1\end{bmatrix}_{\tilde{q}}(-1)^{n-1}\tilde{q}^{(\alpha+\beta)(n-1)+\binom{n}{2}} = \widehat{h}_{\alpha+\beta}(t,a),
\end{aligned}
\tag{4.4}
$$

where we have used (2.30) on the second line and (2.17) on the last line. □

Corollary 4.2. *Let $\alpha \in \mathbb{R}^+$, $\beta \in \mathbb{R}$, $t \in \widetilde{\mathbb{T}}_{(q,h)}^{\sigma^{m+1}(a)}$, where $m \in \mathbb{Z}^+$ is satisfying $m-1 < \alpha \leq m$. Then*

$$
{}_a\nabla_{(q,h)}^{\alpha}\widehat{h}_{\beta}(t,a) = \begin{cases} \widehat{h}_{\beta-\alpha}(t,a), & \beta-\alpha \notin \{-1,\ldots,-m\}, \\ 0, & \beta-\alpha \in \{-1,\ldots,-m\}. \end{cases}
\tag{4.5}
$$

Proof. Proposition 4.1 implies that

$$
{}_a\nabla_{(q,h)}^{\alpha}\widehat{h}_{\beta}(t,a) = \nabla_{(q,h)}^{m}\left({}_a\nabla_{(q,h)}^{-(m-\alpha)}\widehat{h}_{\beta}(t,a)\right) = \nabla_{(q,h)}^{m}\widehat{h}_{m+\beta-\alpha}(t,a).
\tag{4.6}
$$

Then the statement is an immediate consequence of Lemma 2.3. □

Now we are in a position to introduce a (q,h)-discrete analogue of the Mittag-Leffler function. We recall that this function is essentially a generalized exponential function, and its two-parameter form (more convenient in the fractional calculus) can be introduced for $\mathbb{T} = \mathbb{R}$ by the series expansion

$$
E_{\alpha,\beta}(t) = \sum_{k=0}^{\infty}\frac{t^k}{\Gamma(\alpha k+\beta)}, \quad \alpha,\ \beta \in \mathbb{R}^+,\ t \in \mathbb{R}.
\tag{4.7}
$$

The fractional calculus frequently employs (4.7), because the function

$$
t^{\beta-1}E_{\alpha,\beta}(\lambda t^{\alpha}) = \sum_{k=0}^{\infty}\lambda^k\frac{t^{\alpha k+\beta-1}}{\Gamma(\alpha k+\beta)}
\tag{4.8}
$$

(a modified Mittag-Leffler function, see [15]) satisfies under special choices of β a continuous (differential) analogy of (4.1). Some extensions of the definition formula (4.7) and their utilization in special fractional calculus operators can be found in [18, 19].

Considering the discrete calculus, the form (4.8) seems to be much more convenient for discrete extensions than the form (4.7), which requires, among others, the validity of the law

to Contents, page 1

of exponents. The following introduction extends the discrete Mittag-Leffler function defined and studied in [20] for the case $q = h = 1$.

Definition 4.3. Let $\alpha,\ \beta,\ \lambda \in \mathbb{R}$. We introduce the (q,h)-Mittag-Leffler function $E^{s,\lambda}_{\alpha,\beta}(t)$ by the series expansion

$$E^{s,\lambda}_{\alpha,\beta}(t) = \sum_{k=0}^{\infty}\lambda^k \widehat{h}_{\alpha k+\beta-1}(t,s)\left(= \sum_{k=0}^{\infty}\lambda^k \frac{(t-s)^{(\alpha k+\beta-1)}_{(\widetilde{q},h)}}{\Gamma_{\widetilde{q}}(\alpha k+\beta)}\right), \quad s,t \in \widetilde{\mathbb{T}}^{a}_{(q,h)},\ t \geq s\ . \tag{4.9}$$

It is easy to check that the series on the right-hand side converges (absolutely) if $|\lambda|(\nu(t))^{\alpha} < 1$. As it might be expected, the particular (q,h)-Mittag-Leffler function

$$E^{a,\lambda}_{1,1}(t) = \prod_{k=0}^{n-1}\frac{1}{1-\lambda\nu(\rho^k(t))}, \tag{4.10}$$

where $n \in \mathbb{Z}^+$ satisfies $t = \sigma^n(a)$, is a solution of the equation

$$\nabla_{(q,h)}y(t) = \lambda y(t), \qquad t \in \widetilde{\mathbb{T}}^{\sigma(a)}_{(q,h)}, \tag{4.11}$$

that is, it is a discrete (q,h)-analogue of the exponential function.

The main properties of the (q,h)-Mittag-Leffler function are described by the following assertion.

Theorem 4.4. (i) *Let $\eta \in \mathbb{R}^+$ and $t \in \widetilde{\mathbb{T}}^{\sigma(a)}_{(q,h)}$. Then*

$$_a\nabla^{-\eta}_{(q,h)} E^{a,\lambda}_{\alpha,\beta}(t) = E^{a,\lambda}_{\alpha,\beta+\eta}(t). \tag{4.12}$$

(ii) *Let $\eta \in \mathbb{R}^+$, $m \in \mathbb{Z}^+$ be such that $m-1 < \eta \leq m$ and let $\alpha k + \beta - 1 \notin \{0,-1,\ldots,-m+1\}$ for all $k \in \mathbb{Z}^+$. If $t \in \widetilde{\mathbb{T}}^{\sigma^{m+1}(a)}_{(q,h)}$, then*

$$_a\nabla^{\eta}_{(q,h)} E^{a,\lambda}_{\alpha,\beta}(t) = \begin{cases} E^{a,\lambda}_{\alpha,\beta-\eta}(t), & \beta-\eta \notin \{0,-1,\ldots,-m+1\}, \\ \lambda E^{a,\lambda}_{\alpha,\beta-\eta+\alpha}(t), & \beta-\eta \in \{0,-1,\ldots,-m+1\}. \end{cases} \tag{4.13}$$

Proof. The part (i) follows immediately from Proposition 4.1. Considering the part (ii), we can write

$$_a\nabla^{\eta}_{(q,h)} E^{a,\lambda}_{\alpha,\beta}(t) = {}_a\nabla^{\eta}_{(q,h)} \sum_{k=0}^{\infty}\lambda^k \widehat{h}_{\alpha k+\beta-1}(t,a) = \sum_{k=0}^{\infty}\lambda^k\, {}_a\nabla^{\eta}_{(q,h)} \widehat{h}_{\alpha k+\beta-1}(t,a) \tag{4.14}$$

due to the absolute convergence property.

to Contents, page 1

If $k \in \mathbb{Z}^+$, then Corollary 4.2 implies the relation

$$ {}_a\nabla^{\eta}_{(q,h)} \widehat{h}_{\alpha k+\beta-1}(t,a) = \widehat{h}_{\alpha k+\beta-\eta-1}(t,a) \tag{4.15}$$

due to the assumption $\alpha k + \beta - 1 \notin \{0, -1, \ldots, -m+1\}$. If $k = 0$, then two possibilities may occur. If $\beta - \eta \notin \{0, -1, \ldots, -m+1\}$, we get (4.15) with $k = 0$ which implies the validity of $(4.13)_1$. If $\beta - \eta \in \{0, -1, \ldots, -m+1\}$, the nabla (q,h)-fractional derivative of this term is zero and by shifting the index k, we obtain $(4.13)_2$. □

Corollary 4.5. *Let $\alpha \in \mathbb{R}^+$ and $m \in \mathbb{Z}^+$ be such that $m-1 < \alpha \le m$. Then the functions*

$$E^{\alpha,\lambda}_{\alpha,\beta}(t), \qquad \beta = \alpha - m + 1, \ldots, \alpha - 1, \alpha \tag{4.16}$$

define eigenfunctions of the operator ${}_a\nabla^{\alpha}_{(q,h)}$ on each set $[\sigma(a), b] \cap \widetilde{\mathbb{T}}^{\sigma(a)}_{(q,h)}$, where $b \in \widetilde{\mathbb{T}}^{\sigma(a)}_{(q,h)}$ is satisfying $|\lambda|(\nu(b))^{\alpha} < 1$.

Proof. The assertion follows from Theorem 4.4 by the use of $\eta = \alpha$. □

Our final aim is to show that any solution of (4.1) can be written as a linear combination of (q,h)-Mittag-Leffler functions (4.16).

Lemma 4.6. *Let $\alpha \in \mathbb{R}^+$ and $m \in \mathbb{Z}^+$ be such that $m-1 < \alpha \le m$. Then*

$$W_{m-\alpha}\left(E^{\alpha,\lambda}_{\alpha,\alpha-m+1}, E^{\alpha,\lambda}_{\alpha,\alpha-m+2}, \ldots, E^{\alpha,\lambda}_{\alpha,\alpha}\right)(\sigma^m(a)) = \prod_{k=1}^{m} \frac{1}{1-\lambda\left(\nu(\sigma^k(a))\right)^{\alpha}} \neq 0. \tag{4.17}$$

Proof. The case $m = 1$ is trivial. For $m \ge 2$, we can formally write $\lambda E^{\alpha,\lambda}_{\alpha,\alpha-\ell}(t) = E^{\alpha,\lambda}_{\alpha,-\ell}(t)$ for all $t \in \widetilde{\mathbb{T}}^{\sigma^m(a)}_{(q,h)}$ $(\ell = 0, \ldots, m-2)$. Consequently, applying Theorem 4.4, the Wronskian can be expressed as

$$W_{m-\alpha}\left(E^{\alpha,\lambda}_{\alpha,\alpha-m+1}, E^{\alpha,\lambda}_{\alpha,\alpha-m+2}, \ldots, E^{\alpha,\lambda}_{\alpha,\alpha}\right)(\sigma^m(a)) = \det M_m(\sigma^m(a)), \tag{4.18}$$

where

$$M_m(\sigma^m(a)) = \begin{pmatrix} E^{\alpha,\lambda}_{\alpha,1}(\sigma^m(a)) & E^{\alpha,\lambda}_{\alpha,2}(\sigma^m(a)) & \ldots & E^{\alpha,\lambda}_{\alpha,m}(\sigma^m(a)) \\ E^{\alpha,\lambda}_{\alpha,0}(\sigma^m(a)) & E^{\alpha,\lambda}_{\alpha,1}(\sigma^m(a)) & \ldots & E^{\alpha,\lambda}_{\alpha,m-1}(\sigma^m(a)) \\ \ldots & \ldots & \ddots & \ldots \\ E^{\alpha,\lambda}_{\alpha,2-m}(\sigma^m(a)) & E^{\alpha,\lambda}_{\alpha,3-m}(\sigma^m(a)) & \ldots & E^{\alpha,\lambda}_{\alpha,1}(\sigma^m(a)) \end{pmatrix}. \tag{4.19}$$

Using the q-Pascal rule (2.15), we obtain the equality

$$E^{\alpha,\lambda}_{\alpha,i}(\sigma^m(a)) - \nu(\sigma(a)) E^{\alpha,\lambda}_{\alpha,i-1}(\sigma^m(a)) = E^{\sigma(a),\lambda}_{\alpha,i}(\sigma^m(a)), \quad i \in \mathbb{Z},\ i \ge 3-m. \tag{4.20}$$

to Contents, page 1

Starting with the first row, $\binom{m}{2}$ elementary row operations of the type (4.20) transform the matrix $M_m(\sigma^m(a))$ into the matrix

$$\widehat{M}_m(\sigma^m(a)) = \begin{pmatrix} E_{\alpha,1}^{\sigma^{m-1}(a),\lambda}(\sigma^m(a)) & E_{\alpha,2}^{\sigma^{m-1}(a),\lambda}(\sigma^m(a)) & \dots & E_{\alpha,m}^{\sigma^{m-1}(a),\lambda}(\sigma^m(a)) \\ E_{\alpha,0}^{\sigma^{m-2}(a),\lambda}(\sigma^m(a)) & E_{\alpha,1}^{\sigma^{m-2}(a),\lambda}(\sigma^m(a)) & \dots & E_{\alpha,m-1}^{\sigma^{m-2}(a),\lambda}(\sigma^m(a)) \\ \dots & \dots & \ddots & \dots \\ E_{\alpha,2-m}^{a,\lambda}(\sigma^m(a)) & E_{\alpha,3-m}^{a,\lambda}(\sigma^m(a)) & \dots & E_{\alpha,1}^{a,\lambda}(\sigma^m(a)) \end{pmatrix} \tag{4.21}$$

with the property $\det \widehat{M}_m(\sigma^m(a)) = \det M_m(\sigma^m(a))$. By Lemma 2.3, we have

$$\begin{gathered} E_{\alpha,p}^{\sigma^i(a),\lambda}(\sigma^m(a)) - \nu(\sigma^m(a)) E_{\alpha,p-1}^{\sigma^i(a),\lambda}(\sigma^m(a)) = E_{\alpha,p}^{\sigma^i(a),\lambda}\left(\sigma^{m-1}(a)\right), \quad i = 0,\dots,m-2, \\ E_{\alpha,p}^{\sigma^i(a),\lambda}(\sigma^m(a)) - \nu(\sigma^m(a)) E_{\alpha,p-1}^{\sigma^i(a),\lambda}(\sigma^m(a)) = 0, \quad i = m-1, \end{gathered} \tag{4.22}$$

where $p \in \mathbb{Z}$, $p \geq 3 - m + i$. Starting with the last column, using $m-1$ elementary column operations of the type (4.22), we obtain the matrix

$$\left(\begin{array}{c|ccc} E_{\alpha,1}^{\sigma^{m-1}(a),\lambda}(\sigma^m(a)) & 0 & \cdots & 0 \\ \hline E_{\alpha,0}^{\sigma^{m-2}(a),\lambda}(\sigma^m(a)) & & & \\ \vdots & & \widehat{M}_{m-1}(\sigma^{m-1}(a)) & \\ E_{\alpha,2-m}^{a,\lambda}(\sigma^m(a)) & & & \end{array}\right) \tag{4.23}$$

preserving the value of $\det \widehat{M}_m(\sigma^m(a))$. Since

$$E_{\alpha,1}^{\sigma^{m-1}(a),\lambda}(\sigma^m(a)) = \sum_{k=0}^{\infty} \lambda^k (\nu(\sigma^m(a)))^{\alpha k} = \frac{1}{1 - \lambda(\nu(\sigma^m(a)))^{\alpha}}, \tag{4.24}$$

we can observe the recurrence

$$\det \widehat{M}_m(\sigma^m(a)) = \frac{1}{1 - \lambda(\sigma^m(a))^{\alpha}} \det \widehat{M}_{m-1}\left(\sigma^{m-1}(a)\right), \tag{4.25}$$

which implies the assertion. □

Now we summarize the results of Theorem 3.9, Corollary 4.5, and Lemma 4.6 to obtain

Theorem 4.7. *Let $y(t)$ be any solution of (4.1) defined on $[\sigma(a), b] \cap \widetilde{\mathbb{T}}_{(q,h)}^{\sigma(a)}$, where $b \in \widetilde{\mathbb{T}}_{(q,h)}^{\sigma(a)}$ is satisfying $|\lambda|(\nu(b))^{\alpha} < 1$. Then*

to Contents, page 1

$$y(t) = \sum_{j=1}^{m} c_j E_{\alpha,\alpha-m+j}^{a,\lambda}(t), \tag{4.26}$$

where $c_1, \ldots, c_m$ are real constants.

We conclude this paper by the illustrating example.

Example 4.8. Consider the initial value problem

$$\begin{gathered}
{}_a\nabla_{(q,h)}^{\alpha} y(t) = \lambda\, y(t), \quad \sigma^3(a) \le t \le \sigma^n(a),\ 1 < \alpha \le 2, \\
{}_a\nabla_{(q,h)}^{\alpha-1} y(t)\Big|_{t=\sigma^2(a)} = y_{\alpha-1}, \\
{}_a\nabla_{(q,h)}^{\alpha-2} y(t)\Big|_{t=\sigma^2(a)} = y_{\alpha-2},
\end{gathered} \tag{4.27}$$

where n is a positive integer given by the condition $|\lambda|\nu(\sigma^n(a))^{\alpha} < 1$. By Theorem 4.7, its solution can be expressed as a linear combination

$$y(t) = c_1 E_{\alpha,\alpha-1}^{a,\lambda}(t) + c_2 E_{\alpha,\alpha}^{a,\lambda}(t). \tag{4.28}$$

The constants c_1, c_2 can be determined from the system

$$V_{2-\alpha}\left(E_{\alpha,\alpha-1}^{a,\lambda}, E_{\alpha,\alpha}^{a,\lambda}\right)\left(\sigma^2(a)\right) \cdot \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} y_{\alpha-2} \\ y_{\alpha-1} \end{pmatrix} \tag{4.29}$$

with the matrix elements

$$\begin{aligned}
v_{11} = v_{22} &= \frac{[1]_q + \left([\alpha]_q - [1]_q\right)\lambda\nu(\sigma(a))^{\alpha}}{\left(1-\lambda\nu(\sigma(a))^{\alpha}\right)\left(1-\lambda\nu(\sigma^2(a))^{\alpha}\right)}, \\
v_{12} &= \frac{[2]_q\nu(\sigma(a)) + \left([\alpha]_q - [2]_q\right)\lambda\nu(\sigma(a))^{\alpha+1}}{\left(1-\lambda\nu(\sigma(a))^{\alpha}\right)\left(1-\lambda\nu(\sigma^2(a))^{\alpha}\right)}, \\
v_{21} &= \frac{[\alpha]_q\lambda\nu(\sigma(a))^{\alpha-1}}{\left(1-\lambda\nu(\sigma(a))^{\alpha}\right)\left(1-\lambda\nu(\sigma^2(a))^{\alpha}\right)}.
\end{aligned} \tag{4.30}$$

By Lemma 4.6, the matrix $V_{2-\alpha}(E_{\alpha,\alpha-1}^{a,\lambda}, E_{\alpha,\alpha}^{a,\lambda})(\sigma^2(a))$ has a nonzero determinant, hence applying the Cramer rule, we get

$$\begin{aligned}
c_1 &= \frac{y_{\alpha-2}v_{22} - y_{\alpha-1}v_{12}}{W_{2-\alpha}\left(E_{\alpha,\alpha-1}^{a,\lambda}, E_{\alpha,\alpha}^{a,\lambda}\right)(\sigma^2(a))}, \\
c_2 &= \frac{y_{\alpha-1}v_{11} - y_{\alpha-2}v_{21}}{W_{2-\alpha}\left(E_{\alpha,\alpha-1}^{a,\lambda}, E_{\alpha,\alpha}^{a,\lambda}\right)(\sigma^2(a))}.
\end{aligned} \tag{4.31}$$

to Contents, page 1

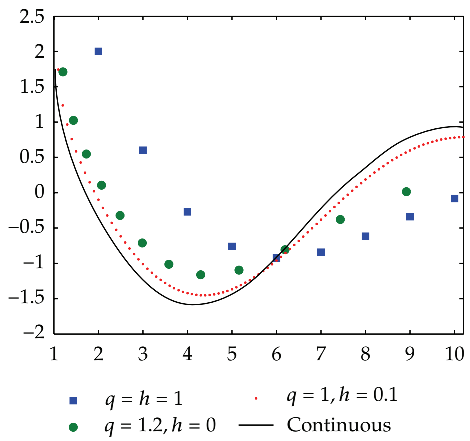


Figure 1: $\alpha = 1.8, a = 1, \lambda = -1/3, y_{\alpha-1} = -1, y_{\alpha-2} = 1$.

Now we make a particular choice of the parameters α, a, λ, $y_{\alpha-1}$ and $y_{\alpha-2}$ and consider the initial value problem in the form

$$
\begin{aligned}
{}_1\nabla^{1.8}_{(q,h)} y(t) &= -\frac{1}{3}\, y(t), \quad \sigma^3(1) \le t \le \sigma^n(1), \\
{}_1\nabla^{0.8}_{(q,h)} y(t)\Big|_{t=\sigma^2(1)} &= -1, \\
{}_1\nabla^{-0.2}_{(q,h)} y(t)\Big|_{t=\sigma^2(1)} &= 1,
\end{aligned}
\tag{4.32}
$$

where n is a positive integer satisfying $\nu(\sigma^n(1)) < 3^{5/9}$. If we take the time scale of integers (the case $q = h = 1$), then the solution $y(t)$ of the corresponding initial value problem takes the form

$$
y(t) = \frac{14}{5}\sum_{k=0}^{\infty}\left(-\frac{1}{3}\right)^k \frac{\prod_{j=1}^{t-2}(j + 1.8k - 0.2)}{(t-2)!} - \frac{2}{15}\sum_{k=0}^{\infty}\left(-\frac{1}{3}\right)^k \frac{\prod_{j=1}^{t-2}(j + 1.8k + 0.8)}{(t-2)!}, \quad t = 2, 3, \ldots. \tag{4.33}
$$

Similarly we can determine $y(t)$ for other choices of q and h. For comparative reasons, Figure 1 depicts (in addition to the above case $q = h = 1$) the solution $y(t)$ under particular choices $q = 1.2$, $h = 0$ (the pure q-calculus), $q = 1$, $h = 0.1$ (the pure h-calculus) and also the solution of the corresponding continuous (differential) initial value problem.

Acknowledgments

The research was supported by the research plan MSM 0021630518 "Simulation modelling of mechatronic systems" of the Ministry of Education, Youth and Sports of the Czech Republic, by Grant P201/11/0768 of the Czech Grant Agency and by Grant FSI-J-10-55 of the FME, Brno University of Technology.

to Contents, page 1

References

[1] R. P. Agarwal, "Certain fractional q-integrals and q-derivatives," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 66, pp. 365–370, 1969.

[2] J. B. Diaz and T. J. Osler, "Differences of fractional order," *Mathematics of Computation*, vol. 28, pp. 185–202, 1974.

[3] H. L. Gray and N. F. Zhang, "On a new definition of the fractional difference," *Mathematics of Computation*, vol. 50, no. 182, pp. 513–529, 1988.

[4] K. S. Miller and B. Ross, "Fractional difference calculus," in *Univalent Functions, Fractional Calculus, and Their Applications (Koriyama, 1988)*, Ellis Horwood Series: Mathematics and Its Applications, pp. 139–152, Horwood, Chichester, UK, 1989.

[5] F. M. Atici and P. W. Eloe, "A transform method in discrete fractional calculus," *International Journal of Difference Equations*, vol. 2, no. 2, pp. 165–176, 2007.

[6] F. M. Atici and P. W. Eloe, "Fractional q-calculus on a time scale," *Journal of Nonlinear Mathematical Physics*, vol. 14, no. 3, pp. 341–352, 2007.

[7] J. Čermák and L. Nechvátal, "On (q, h)-analogue of fractional calculus," *Journal of Nonlinear Mathematical Physics*, vol. 17, no. 1, pp. 51–68, 2010.

[8] M. Bohner and A. Peterson, *Dynamic equations on time scales*, Birkhauser, Boston, Mass, USA, 2001.

[9] M. Bohner and A. Peterso, Eds., *Advances in Dynamic Equations on Time Scales*, Birkhauser, Boston, Mass, USA, 2003.

[10] G. A. Anastassiou, "Foundations of nabla fractional calculus on time scales and inequalities," *Computers & Mathematics with Applications*, vol. 59, no. 12, pp. 3750–3762, 2010.

[11] G. E. Andrews, R. Askey, and R. Roy, *Special Functions*, vol. 71 of *Encyclopedia of Mathematics and Its Applications*, Cambridge University Press, Cambridge, UK, 1999.

[12] F. M. Atici and P. W. Eloe, "Initial value problems in discrete fractional calculus," *Proceedings of the American Mathematical Society*, vol. 137, no. 3, pp. 981–989, 2009.

[13] F. M. Atici and P. W. Eloe, "Discrete fractional calculus with the nabla operator," *Electronic Journal of Qualitative Theory of Differential Equations*, vol. 2009, no. 2, p. 12, 2009.

[14] Z. S. I. Mansour, "Linear sequential q-difference equations of fractional order," *Fractional Calculus & Applied Analysis*, vol. 12, no. 2, pp. 159–178, 2009.

[15] I. Podlubny, *Fractional Differential Equations*, vol. 198 of *Mathematics in Science and Engineering*, Academic Press, San Diego, Calif, USA, 1999.

[16] A. Nagai, "On a certain fractional q-difference and its eigen function," *Journal of Nonlinear Mathematical Physics*, vol. 10, supplement 2, pp. 133–142, 2003.

[17] J. Čermák and T. Kisela, "Note on a discretization of a linear fractional differential equation," *Mathematica Bohemica*, vol. 135, no. 2, pp. 179–188, 2010.

[18] A. A. Kilbas, M. Saigo, and R. K. Saxena, "Solution of Volterra integrodifferential equations with generalized Mittag-Leffler function in the kernels," *Journal of Integral Equations and Applications*, vol. 14, no. 4, pp. 377–396, 2002.

[19] A. A. Kilbas, M. Saigo, and R. K. Saxena, "Generalized Mittag-Leffler function and generalized fractional calculus operators," *Integral Transforms and Special Functions*, vol. 15, no. 1, pp. 31–49, 2004.

[20] F. M. Atici and P. W. Eloe, "Linear systems of fractional nabla difference equations," *Rocky Mountain Journal of Mathematics*, vol. 41, pp. 353–370, 2011.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 920412, 14 pages
doi:10.1155/2011/920412

Research Article

Estimates of Exponential Stability for Solutions of Stochastic Control Systems with Delay

Irada A. Dzhalladova,[1] Jaromír Baštinec,[2] Josef Diblík,[2,3] and Denys Y. Khusainov[4]

[1] *Department of Mathematics, Kyiv National Economic University, Peremogy Avenue, Kyiv 03038, Ukraine*
[2] *Department of Mathematics, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technická 8, 616 00 Brno, Czech Republic*
[3] *Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Brno University of Technology, Veveří 331/95, 602 00 Brno, Czech Republic*
[4] *Faculty of Cybernetics, Taras Shevchenko National University of Kyiv, Vladimirskaya Street, Kyiv 01601, Ukraine*

Correspondence should be addressed to Irada A. Dzhalladova, irada-05@mail.ru

Received 2 February 2011; Accepted 28 March 2011

Academic Editor: Miroslava Růžičková

A nonlinear stochastic differential-difference control system with delay of neutral type is considered. Sufficient conditions for the exponential stability are derived by using Lyapunov-Krasovskii functionals of quadratic form with exponential factors. Upper bound estimates for the exponential rate of decay are derived.

1. Introduction

The theory and applications of functional differential equations form an important part of modern nonlinear dynamics. Such equations are natural mathematical models for various real life phenomena where the aftereffects are intrinsic features of their functioning. In recent years, functional differential equations have been used to model processes in different areas such as population dynamics and ecology, physiology and medicine, economics, and other natural sciences [1–3]. In many of the models the initial data and parameters are subjected to random perturbations, or the dynamical systems themselves represent stochastic processes. For this reason, stochastic functional differential equations are widely studied [4, 5].

One of the principal problems of the corresponding mathematical analysis of equations is a comprehensive study of their global dynamics and the related prediction of

to Contents, page 1

long-term behaviors in applied models. Of course, the problem of stability of a particular solution plays a significant role. Therefore, the study of stability of linear equations is the first natural and important step in the analysis of more complex nonlinear systems.

When applying the mathematical theory to real-world problems a mere statement of the stability in the system is hardly sufficient. In addition to stability as such, it is of significant importance to obtain constructive and verifiable estimates of the rate of convergence of solutions in time. One of the principal tools used in the related studies is the second Lyapunov method [6–8]. For functional differential equations, this method has been developing in two main directions in recent years. The first one is the method of finite Lyapunov functions with the additional assumption of Razumikhin type [9, 10]. The second one is the method of Lyapunov-Krasovskii functionals [11, 12]. For stochastic functional differential equations, some aspects of these two lines of research have been developed, for example, in [11, 13–19] and [11, 18, 20–25], respectively. In the present paper, by using the method of Lyapunov-Krasovskii functionals, we derive sufficient conditions for stability together with the rate of convergence to zero of solutions for a class of linear stochastic functional differential equation of a neutral type.

2. Preliminaries

In solving control problems for linear systems, very often, a scalar function $u = u(x)$ needs to be found such that the system

$$\dot{x}(t) = Ax(t) + bu(x(t)) \tag{2.1}$$

is asymptotically stable. Frequently, such a function depends on a scalar argument which is a linear combination of phase coordinates and its graph lies in the first and the third quadrants of the plane. An investigation of the asymptotic stability of systems with a control function

$$u(x(t)) = f(\sigma(t)), \quad \sigma(t) = c^T x(t), \tag{2.2}$$

that is, an investigation of systems

$$\dot{x}(t) = Ax(t) + bf(\sigma(t)), \quad \sigma(t) = c^T x(t), \tag{2.3}$$

with a function f satisfying $f(0) = 0$, $f(\sigma)(k\sigma - f(\sigma)) > 0$ for $\sigma \neq 0$ and a $k > 0$ is called an analysis of the absolute stability of control systems [26]. One of the fundamental methods (called a frequency method) was developed by Gelig et al. (see, e.g., the book [27]). Another basic method is the method of Lyapunov's functions and Lyapunov-Krasovskii functionals. Very often, the appropriate Lyapunov functions and Lyapunov-Krasovskii functionals are constructed as quadratic forms with integral terms containing a given nonlinearity [28, 29]. An overview of the present state can be found, for example, in [30, 31]. Problems of absolute stability of stochastic equations are treated, for example, in [11, 14, 15, 24].

to Contents, page 1

3. Main Results

Consider the following control system of stochastic differential-difference equations of a neutral type

$$\begin{aligned} d[x(t) - Dx(t-\tau)] = {} & \left[A_0 x(t) + A_1 x(t-\tau) + a_2 f(\sigma(t))\right] dt \\ & + \left[B_0 x(t) + B_1 x(t-\tau) + b_2 f(\sigma(t))\right] dw(t), \end{aligned} \tag{3.1}$$

where

$$\sigma(t) := c^T [x(t) - Dx(t-\tau)], \tag{3.2}$$

$x : [0,\infty) \to \mathbb{R}^n$ is an n-dimensional column vector, A_0, A_1, B_0, B_1, and D are real $n \times n$ constant matrices, a_2, b_2, and c are $n \times 1$ constant vectors, $f : \mathbb{R} \to \mathbb{R}$ is a continuous function, $\tau > 0$ is a constant delay, and $w(t)$ is a standard scalar Wiener process with

$$M\{dw(t)\} = 0, \qquad M\left\{dw^2(t)\right\} = dt, \qquad M\{dw(t_1)\, dw(t_2), t_1 \neq t_2\} = 0. \tag{3.3}$$

An F_t-measurable random process $\{x(t) \equiv x(t,\omega)\}$ is called a solution of (3.1) if it satisfies, with a probability one, the following integral equation

$$\begin{aligned} x(t) = {} & Dx(t-\tau) + [x(0) - Dx(-\tau)] \\ & + \int_0^t \left[A_0 x(s) + A_1 x(s-\tau) + a_2 f(\sigma(s))\right] ds \\ & + \int_0^t \left[B_0 x(s) + B_1 x(s-\tau) + b_2 f(\sigma(s))\right] dw(s), \quad t \geq 0 \end{aligned} \tag{3.4}$$

and the initial conditions

$$x(t) = \varphi(t), \quad x'(t) = \psi(t), \quad t \in [-\tau, 0], \tag{3.5}$$

where $\varphi, \psi : [-\tau, 0] \to \mathbb{R}^n$ are continuous functions. Here and in the remaining part of the paper, we will assume that the initial functions φ and ψ are continuous random processes. Under those assumptions, a solution to the initial value problem (3.1), (3.5) exists and is unique for all $t \geq 0$ up to its stochastic equivalent solution on the space (Ω, F, P) [4].

to Contents, page 1

We will use the following norms of matrices and vectors

$$\begin{aligned}
\|A\| &:= \sqrt{\lambda_{\max}(A^T A)}\ , \\
\|x(t)\| &:= \sqrt{\sum_{i=1}^{n} x_i^2(t)}, \\
\|x(t)\|_\tau &:= \max_{-\tau\le s\le 0}\{\|x(t+s)\|\}, \\
\|x(t)\|_{\tau,\gamma}^2 &:= \int_{t-\tau}^{t} e^{-\gamma(t-s)}\|x(s)\|^2 ds,
\end{aligned} \tag{3.6}$$

where $\lambda_{\max}(*)$ is the largest eigenvalue of the given symmetric matrix (similarly, the symbol $\lambda_{\min}(*)$ denotes the smallest eigenvalue of the given symmetric matrix), and γ is a positive parameter.

Throughout this paper, we assume that the function f satisfies the inequality

$$0 \le f(\sigma)\sigma \le k\sigma^2 \tag{3.7}$$

if $\sigma \in \mathbb{R}$ where k is a positive constant.

For the reader's convenience, we recall that the zero solution of (3.1) is called stable in the square mean if, for every $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon) > 0$ such that every solution $x = x(t)$ of (3.1) satisfies $M\{\|x(t)\|^2\} < \varepsilon$ provided that the initial conditions (3.5) are such that $\|\varphi(0)\|_\tau < \delta$ and $\|\psi(0)\|_\tau < \delta$. If the zero solution is stable in the square mean and, moreover,

$$\lim_{t\to+\infty} M\left\{\|x(t)\|^2\right\} = 0, \tag{3.8}$$

then it is called asymptotically stable in the square mean.

Definition 3.1. If there exist positive constants N, γ, and θ such that the inequality

$$M\left\{\|x(t)\|_{\tau,\gamma}^2\right\} \le N\ \|x(0)\|_\tau^2\ e^{-\theta t} \tag{3.9}$$

holds on $[0,\infty)$, then the zero solution of (3.1) is called exponentially γ-integrally stable in the square mean.

In this paper, we prove the exponential γ-integral stability in the square mean of the differential-difference equation with constant delay (3.1). We employ the method of stochastic Lyapunov-Krasovskii functionals. In [11, 18, 22, 24] the Lyapunov-Krasovskii functional is chosen to be of the form

$$V[x(t),t] = h[x(t) - cx(t-\tau)]^2 + g\int_{-\tau}^{0} x^2(t+s)ds, \tag{3.10}$$

to Contents, page 1

where constants $h > 0$ and $g > 0$ are such that the total stochastic differential of the functional along solutions is negative definite.

In the present paper, we consider the Lyapunov-Krasovskii functional in the following form:

$$\begin{aligned} V[x(t),t] &= [x(t) - Dx(t-\tau)]^T H[x(t) - Dx(t-\tau)] \\ &\quad + \int_{t-\tau}^{t} e^{-\gamma(t-s)} x^T(s) G x(s) ds + \beta \int_0^{\sigma(t)} f(\xi) d\xi, \end{aligned} \tag{3.11}$$

where constants $\gamma > 0$, $\beta > 0$ and $n \times n$ positive definite symmetric matrices G, H are to be restricted later on. This allows us not only to derive sufficient conditions for the stability of the zero solution but also to obtain coefficient estimates of the rate of the exponential decay of solutions.

We set

$$P := \begin{pmatrix} H & -HD \\ -D^T H & D^T H D \end{pmatrix}. \tag{3.12}$$

Then, by using introduced norms, the functional (3.11) yields two-sided estimates

$$\begin{aligned} \lambda_{\min}(G)\|x(t)\|_{\tau,\gamma}^2 \le V[x(t),t] &\le \left[\lambda_{\max}(P) + 0.5\beta k\|c\|^2\right]\|x(t)\|^2 \\ &\quad + \left[\lambda_{\max}(P) + 0.5\beta k\left\|c^T D\right\|^2\right]\|x(t-\tau)\|^2 + \lambda_{\max}(G)\|x(t)\|_{\tau,\gamma}^2, \end{aligned} \tag{3.13}$$

where $t \in [0, \infty)$.

We will use an auxiliary $(2n+1) \times (2n+1)$-dimensional matrix:

$$S = S(\beta, \gamma, \nu, G, H) := \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix}, \tag{3.14}$$

where

$$
\begin{aligned}
s_{11} &:= -A_0H - HA_0 - B_0^T HB_0 - G, \\
s_{12} &:= A_0^T HD - HA_1 - B_0^T HB_1, \\
s_{13} &:= -Ha_2 - B_0^T Hb_2 - \frac{1}{2}(\beta A_0 + \nu I)^T c, \\
s_{21} &:= s_{12}^T, \\
s_{22} &:= D^T HA_1 + A_1^T HD - B_1^T HB_1 + e^{-\gamma\tau} G, \\
s_{23} &:= D^T Ha_2 - B_1^T Hb_2 - \frac{1}{2}\beta A_1 c, \\
s_{31} &:= s_{13}^T, \\
s_{32} &:= s_{23}^T, \\
s_{33} &:= \frac{\nu}{k} - b_2^T Hb_2 - \beta c^T a_2,
\end{aligned}
\tag{3.15}
$$

where ν is a parameter.

Now we establish our main result on the exponential γ-integral stability of a trivial solution in the square mean of system (3.1) when $t \to \infty$.

Theorem 3.2. *Let $\|D\| < 1$. Let there exist positive constants β, γ, ν and positive definite symmetric matrices G, H such that the matrix S is positively definite as well. Then the zero solution of the system (3.1) is exponentially γ-integrally stable in the square mean on $[0, \infty)$. Moreover, every solution $x(t)$ of (3.1) satisfies the inequality*

$$
M\left\{\|x(t)\|_{\tau,\gamma}^2\right\} \le N\|x(0)\|_\tau^2 \; e^{-\theta t}
\tag{3.16}
$$

for all $t \ge 0$ where

$$
\begin{aligned}
N &:= \frac{1}{\lambda_{\min}(G)} \cdot \left(2\lambda_{\max}(P) + 0.5\beta k\|c\|^2 + 0.5\beta k\left\|c^T D\right\|^2 + \frac{1}{\gamma}\lambda_{\max}(G)\right), \\
\theta &:= \min\left\{\frac{\gamma\lambda_{\min}(G)}{\lambda_{\max}(G)}, \frac{\lambda_{\min}(S)}{\lambda_{\max}(P) + 0.5\beta k\|c\|^2}\right\}.
\end{aligned}
\tag{3.17}
$$

to Contents, page 1

Proof. We will apply the method of Lyapunov-Krasovskii functionals using functional (3.11). Using the Itô formula, we compute the stochastic differential of (3.11) as follows

$$\begin{aligned}
dV[x(t),t] = \Big(&\left[A_0x(t)+A_1x(t-\tau)+a_2f(\sigma(t))\right]^T dt \\
&+\left[B_0x(t)+B_1x(t-\tau)+b_2f(\sigma(t))\right]^T dw(t)\Big) \\
&\times H[x(t)-Dx(t-\tau)]+[x(t)-Dx(t-\tau)]^T \\
&\times H\Big(\left[A_0x(t)+A_1x(t-\tau)+a_2f(\sigma(t))\right]dt \\
&\qquad+\left[B_0x(t)+B_1x(t-\tau)+b_2f(\sigma(t))\right]^T dw(t)\Big) \\
&+\left[B_0x(t)+B_1x(t-\tau)+b_2f(\sigma(t))\right]^T \\
&\times H\left[B_0x(t)+B_1x(t-\tau)+b_2f(\sigma(t))\right]d\Big(w^2(t)\Big) \\
&+x^T(t)Gx(t)dt-e^{-\gamma\tau}x^T(t-\tau)Gx(t-\tau)dt+\beta f(\sigma(t))c^T \\
&\times\Big(\left[A_0x(t)+A_1x(t-\tau)+a_2f(\sigma(t))\right]dt \\
&\quad+\left[B_0x(t)+B_1x(t-\tau)+b_2f(\sigma(t))\right]^T dw(t)\Big) \\
&-\gamma\int_{t-\tau}^{t}e^{-\gamma(t-s)}x^T(s)Gx(s)ds\,dt.
\end{aligned} \tag{3.18}$$

Taking the mathematical expectation we obtain (we use properties (3.3))

$$\begin{aligned}
M\{dV[x(t),t]\} = M\Big\{&\left[A_0x(t)+A_1x(t-\tau)+a_2f(\sigma(t))\right]^T \\
&\times H[x(t)-Dx(t-\tau)]dt\Big\} \\
&+M\Big\{[x(t)-Dx(t-\tau)]^T \\
&\quad\times H\left[A_0x(t)+A_1x(t-\tau)+a_2f(\sigma(t))\right]dt\Big\} \\
&+M\Big\{\left[B_0x(t)+B_1x(t-\tau)+b_2f(\sigma(t))\right]^T \\
&\quad\times H\left[B_0x(t)+B_1x(t-\tau)+b_2f(\sigma(t))\right]d\Big(w^2(t)\Big)\Big\} \\
&+M\Big\{\Big[x^T(t)Gx(t)dt-e^{-\gamma\tau}x^T(t-\tau)Gx(t-\tau)dt\Big]\Big\} \\
&+\beta M\Big\{f(\sigma(t))c^T\left[A_0x(t)+A_1x(t-\tau)+a_2f(\sigma(t))\right]dt\Big\} \\
&-\gamma M\left\{\int_{t-\tau}^{t}e^{-\gamma(t-s)}x^T(s)Gx(s)ds\,dt\right\}.
\end{aligned} \tag{3.19}$$

to Contents, page 1

Utilizing the matrix S defined by (3.14), the last expression can be rewritten in the following vector matrix form

$$\frac{d}{dt}M\{V[x(t),t]\} = -M\left\{\left(x^T(t), x^T(t-\tau), f(\sigma(t))\right) \times S \times \left(x^T(t), x^T(t-\tau), f(\sigma(t))\right)^T\right\}$$
$$- \nu\left[\sigma(t) - \frac{f(\sigma(t))}{k}\right]f(\sigma(t)) - \gamma M\left\{\int_{t-\tau}^{t} e^{-\gamma(t-s)}x^T(s)Gx(s)ds\right\}. \tag{3.20}$$

We will show next that solutions of (3.1) decay exponentially by calculating the corresponding exponential rate.

The full derivative of the mathematical expectation for the Lyapunov-Krasovskii functional (3.11) satisfies

$$\begin{aligned}\frac{d}{dt}M\{V[x(t),t]\} &\le -\lambda_{\min}(S)M\left\{\|x(t)\|^2\right\}\\ &\quad - \lambda_{\min}(S)M\left\{\|x(t-\tau)\|^2\right\}\\ &\quad - \gamma\lambda_{\min}(G)M\left\{\|x(t)\|^2_{\tau,\gamma}\right\}.\end{aligned} \tag{3.21}$$

In the following we will use inequalities being a consequence of (3.13).

$$\begin{aligned}\lambda_{\min}(G)M\left\{\|x(t)\|^2_{\tau,\gamma}\right\} &\le M\{V[x(t)]\}\\ &\le \left[\lambda_{\max}(P) + 0.5\beta k\|c\|^2\right] \times M\left\{\|x(t)\|^2\right\}\\ &\quad + \left[\lambda_{\max}(P) + 0.5\beta k\left\|c^T D\right\|^2\right]M\left\{\|x(t-\tau)\|^2\right\}\\ &\quad + \lambda_{\max}(G)M\left\{\|x(t)\|^2_{\tau,\gamma}\right\}.\end{aligned} \tag{3.22}$$

Let us derive conditions for the coefficients of (3.1) and parameters of the Lyapunov-Krasovskii functional (3.11) such that the following inequality:

$$\frac{d}{dt}M\{V[x(t),t]\} \le -\theta M\{V[x(t),t]\} \tag{3.23}$$

holds. We use a sequence of the following calculations supposing that either inequality

$$\gamma\lambda_{\min}(G) - \frac{\lambda_{\min}(S)}{\lambda_{\max}(P) + 0.5\beta k|c|^2}\lambda_{\max}(G) \ge 0 \tag{3.24}$$

to Contents, page 1

holds, or the opposite inequality

$$\gamma\lambda_{\min}(G) - \frac{\lambda_{\min}(S)}{\lambda_{\max}(P) + 0.5\beta k|c|^2}\lambda_{\max}(G) \leq 0 \tag{3.25}$$

is valid.

(1) Let inequality (3.24) holds. Rewrite the right-hand part of inequality (3.22) in the form

$$\begin{aligned} -M\left\{\|x(t)\|^2\right\} \leq{}& \frac{1}{\lambda_{\max}(P) + 0.5\beta k\|c\|^2} \\ &\times\left[-M\{V[x(t),t]\} + \lambda_{\max}(G)M\left\{\|x(t)\|^2_{\tau,\gamma}\right\}\right. \\ &\left.+\left[\lambda_{\max}(P) + 0.5\beta k\left\|c^T D\right\|^2\right]M\left\{\|x(t-\tau)\|^2\right\}\right] \end{aligned} \tag{3.26}$$

and substitute the latter into inequality (3.21). This results in

$$\begin{aligned} \frac{d}{dt}M\{V[x(t),t]\} \leq{}& -\frac{\lambda_{\min}(S)}{\lambda_{\max}(P) + 0.5\beta k\|c\|^2} \\ &\times\left[-M\{V[x(t),t]\} + \lambda_{\max}(G)M\left\{\|x(t)\|^2_{\tau,\gamma}\right\}\right. \\ &\left.+\left[\lambda_{\max}(P) + 0.5\beta k\left\|c^T D\right\|^2\right]M\left\{\|x(t-\tau)\|^2\right\}\right] \\ &-\gamma\lambda_{\min}(G)M\left\{\|x(t)\|^2_{\tau,\gamma}\right\} - \lambda_{\min}(S)M\left\{\|x(t-\tau)\|^2\right\}, \end{aligned} \tag{3.27}$$

or, equivalently,

$$\begin{aligned} \frac{d}{dt}M\{V[x(t),t]\} \leq{}& -\frac{\lambda_{\min}(S)}{\lambda_{\max}(P) + 0.5\beta k\|c\|^2}M\{V[x(t),t]\} \\ &-\lambda_{\min}(S)\left(1 - \frac{\lambda_{\max}(P) + 0.5\beta k\left\|c^T D\right\|^2}{\lambda_{\max}(P) + 0.5\beta k\|c\|^2}\right)M\left\{\|x(t-\tau)\|^2\right\} \\ &-\left(\gamma\lambda_{\min}(G) - \frac{\lambda_{\min}(S)}{\lambda_{\max}(P) + 0.5\beta k\|c\|^2}\lambda_{\max}(G)\right)M\left\{\|x(t)\|^2_{\tau,\gamma}\right\}. \end{aligned} \tag{3.28}$$

The inequality

$$\frac{\lambda_{\max}(P) + 0.5\beta k\left\|c^T D\right\|^2}{\lambda_{\max}(P) + 0.5\beta k\|c\|^2} \leq 1 \tag{3.29}$$

to Contents, page 1

always holds. Because inequality (3.24) is valid, a differential inequality

$$\begin{aligned}\frac{d}{dt}M\{V[x(t),t]\} &\le -\frac{\lambda_{\min}(S)}{\lambda_{\max}(P)+0.5\beta k\|c\|^2}M\{V[x(t),t]\} \\ &\le -\theta M\{V[x(t),t]\}\end{aligned} \tag{3.30}$$

will be true as well.

(2) Let inequality (3.25) hold. We rewrite the right-hand side of inequality (3.22) in the form

$$\begin{aligned}-M\left\{\|x(t)\|^2_{\tau,\gamma}\right\} \le \frac{1}{\lambda_{\max}(G)} \times \Bigg(&-M\{V[x(t),t]\} + \left(\lambda_{\max}(P)+0.5\beta k\|c\|^2\right)M\left\{\|x(t)\|^2\right\} \\ &+\left[\lambda_{\max}(P)+0.5\beta k\left\|c^TD\right\|^2\right]M\left\{\|x(t-\tau)\|^2\right\}\Bigg)\end{aligned} \tag{3.31}$$

and substitute the latter again into inequality (3.21). This results in

$$\begin{aligned}\frac{d}{dt}M\{V[x(t),t]\} \le -\lambda_{\min}(S)M\left\{\|x(t)\|^2\right\} - \lambda_{\min}(S)M\left\{\|x(t-\tau)\|^2\right\} + \gamma\frac{\lambda_{\min}(G)}{\lambda_{\max}(G)} \\ \times\left\{ -M\{V[x(t),t]\} + \left(\lambda_{\max}(P)+0.5\beta k\|c\|^2\right)M\left\{\|x(t)\|^2\right\}\right. \\ \left.+\left[\lambda_{\max}(P)+0.5\beta k\left\|c^TD\right\|^2\right]M\left\{\|x(t-\tau)\|^2\right\}\right\}\end{aligned} \tag{3.32}$$

or in

$$\begin{aligned}\frac{d}{dt}M\{V[x(t),t]\} \le -\gamma\frac{\lambda_{\min}(G)}{\lambda_{\max}(G)}M\{V[x(t),t]\} \\ -\left(\lambda_{\min}(S) - \frac{\lambda_{\max}(P)+0.5\beta k\|c\|^2}{\lambda_{\max}(G)}\gamma\lambda_{\min}(G)\right)M\left\{\|x(t)\|^2\right\} \\ -\left(\lambda_{\min}(S) - \frac{\gamma\lambda_{\min}(G)\left[\lambda_{\max}(P)+0.5\beta k\left\|c^TD\right\|^2\right]}{\lambda_{\max}(G)}\right)M\left\{\|x(t-\tau)\|^2\right\}.\end{aligned} \tag{3.33}$$

Because inequality (3.25) is valid, a differential inequality

$$\frac{d}{dt}M\{V[x(t),t]\} \le -\gamma\frac{\lambda_{\min}(G)}{\lambda_{\max}(G)}M\{V[x(t),t]\} \le -\theta M\{V[x(t),t]\} \tag{3.34}$$

will be valid as well.

to Contents, page 1

Analysing inequalities (3.30) and (3.34) we conclude that (3.23) always holds. Solving inequality (3.23) we obtain

$$M\{V[x(t),t]\} \le M\{V[x(0),0]\}e^{-\theta t}. \tag{3.35}$$

Now we derive estimates of the rate of the exponential decay of solutions. We use inequalities (3.22), (3.35). It is easy to see that

$$\begin{aligned}
\lambda_{\min}(G)M\left\{\|x(t)\|^2_{\tau,\gamma}\right\} &\le M\{V[x(t),t]\} \le M\{V[x(0),0]\}e^{-\theta t} \\
&\le \left(\left(\lambda_{\max}(P) + 0.5\beta k\|c\|^2\right)\|x(0)\|^2\right. \\
&\quad \left. + \left[\lambda_{\max}(P) + 0.5\beta k\left\|c^T D\right\|^2\right]\|x(-\tau)\|^2 + \lambda_{\max}(G)\|x(0)\|^2_{\tau,\gamma}\right)e^{-\theta t} \\
&\le \left(2\lambda_{\max}(P) + 0.5\beta k\|c\|^2 + 0.5\beta k\left\|c^T D\right\|^2 + \frac{1}{\gamma}\lambda_{\max}(G)\right)\|x(0)\|^2_{\tau}e^{-\theta t}.
\end{aligned} \tag{3.36}$$

Now, inequality (3.16) is a simple consequence of the latter chain of inequalities. □

4. A Scalar Case

As an example, we will apply Theorem 3.2 to a scalar control stochastic differential-difference equation of a neutral type

$$\begin{aligned}
d[x(t) - d_0x(t-\tau)] = {}& \left[a_0x(t) + a_1x(t-\tau) + a_2f(\sigma(t))\right]dt \\
&+ \left[b_0x(t) + b_1x(t-\tau) + b_2f(\sigma(t))\right]dw(t),
\end{aligned} \tag{4.1}$$

where $\sigma(t) = c[x(t) - d_0x(t-\tau)]$, $x \in \mathbb{R}$, a_0, a_1, a_2, b_0, b_1, d_2, d_0, and c are real constants, $\tau > 0$ is a constant delay, and $w(t)$ is a standard scalar Wiener process satisfying (3.3). An F_t-measurable random process $\{x(t) \equiv x(t,\omega)\}$ is called a solution of (4.1) if it satisfies, with a probability one, the following integral equation:

$$\begin{aligned}
x(t) = {}& d_0x(t-\tau) + [x(0) - d_0x(-\tau)] \\
&+ \int_0^t \left[a_0x(s) + a_1x(s-\tau) + a_2f(\sigma(t))\right]ds \\
&+ \int_0^t \left[b_0x(s) + b_1x(s-\tau) + b_2f(\sigma(t))\right]dw(s), \quad t \ge 0.
\end{aligned} \tag{4.2}$$

to Contents, page 1

The Lyapunov-Krasovskii functional V reduces to

$$V[x(t), t] = [x(t) - d_0 x(t-\tau)]^2 + g \int_{t-\tau}^{t} e^{-\gamma(t-s)} x^2(s) ds + \beta \int_0^{\sigma(t)} f(\xi) d\xi, \tag{4.3}$$

where we assume $g > 0$ and $\beta > 0$. The matrix S reduces to (for simplicity we set $H = (1)$)

$$S = S(g, \beta, \gamma, \nu) := \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix} \tag{4.4}$$

and has entries

$$\begin{aligned}
s_{11} &:= -2a_0 - b_0^2 - g, \\
s_{12} &:= a_0 d_0 - a_1 - b_0 b_1, \\
s_{13} &:= -a_2 - b_0 b_2 - \frac{1}{2}(\beta a_0 + \nu) c, \\
s_{21} &:= s_{12}, \\
s_{22} &:= 2a_1 d_0 - b_1^2 + e^{-\gamma\tau} g, \\
s_{23} &:= a_2 d_0 - b_1 b_2 - 0.5\beta a_1 c, \\
s_{31} &:= s_{13}, \\
s_{32} &:= s_{23}, \\
s_{33} &:= \frac{\nu}{k} - b_2^2 - \beta c a_2,
\end{aligned} \tag{4.5}$$

where ν is a parameter. Therefore, the above calculation yields the following result.

Theorem 4.1. *Let $|d_0| < 1$. Assume that positive constants β, γ, g, and ν are such that the matrix S is positive definite. Then the zero solution of (4.1) is exponentially γ-integrally stable in the square mean on $[0, \infty)$. Moreover, every solution $x(t)$ satisfies the following convergence estimate:*

$$M\left\{ \|x(t)\|_{\tau,\gamma}^2 \right\} \leq N \|x(0)\|_\tau^2 e^{-\theta t} \tag{4.6}$$

for all $t \geq 0$ where

$$\begin{aligned}
N &:= \frac{1}{g}\left(2 + 2d_0^2 + 0.5\beta k c^2 + 0.5\beta k (c d_0)^2\right) + \frac{1}{\gamma}, \\
\theta &:= \min\left\{ \gamma, \frac{\lambda_{\min}(S)}{1 + d_0^2 + 0.5\beta k c^2} \right\}.
\end{aligned} \tag{4.7}$$

to Contents, page 1

Acknowledgments

The second author was supported by Grant P201/10/1032 of Czech Grant Agency, and by Project FEKT/FSI-S-11-1-1159. The third author was supported by Grant P201/11/0768 of Czech Grant Agency, and by the Council of Czech Government MSM 0021630503. The fourth author was supported by Project M/34-2008 of Ukrainian Ministry of Education, Ukraine and by Grant P201/10/1032 of Czech Grant Agency.

References

[1] L. Glass and M. C. Mackey, *From Clocks to Chaos*, The Rhythms of Life, Princeton University Press, Princeton, NJ, USA, 1988.
[2] V. Kolmanovskii and A. Myshkis, *Introduction to the Theory and Applications of Functional-Differential Equations*, vol. 463 of *Mathematics and Its Applications*, Kluwer Academic, Dordrecht, The Netheralnds, 1999.
[3] Y. Kuang, *Delay Differential Equations with Applications in Population Dynamics*, vol. 191 of *Mathematics in Science and Engineering*, Academic Press, Boston, Mass, USA, 1993.
[4] I. I. Gikhman and A. V. Skorokhod, *Stochastic Differential Equations*, Naukova Dumka, Kiev, Ukraine, 1968.
[5] R. Z. Khasminskii, *Stability of Systems of Differential Equations under Stochastic Perturbations of Parameters*, Science Publishing House, Moscow, Russia, 1969.
[6] A. A. Martynuk and A. A. Martynjuk, *Stability of Motion of Complex Systems*, Naukova Dumka, Kiev, Ukraine, 1975.
[7] E. A. Barbashin, *Lyapunov's Functions*, Nauka, Moskow, Russia, 1970.
[8] E. A. Barbashin, *Introduction to the Theory of Stability*, Wolters-Noordhoff, Groningen, The Netherlands, 1970.
[9] D. Ya. Khusainov and A. V. Shatyrko, *Method of Lyapunov Functions in the Study of Stability Functional Differential Systems*, Publishing Company of Kiev National University, Kiev, Ukraine, 1997.
[10] B. S. Razumikhin, *Stability in Hereditary Systems*, Science Publishing House, Moscow, Russia, 1988.
[11] D. G. Korenevskii, *Stability of Solutions in Deterministic and Stochastic Differential-Difference Equations*, Algebraic Criteria, Naukova Dumka, Kiev, Ukraine, 1992.
[12] N. N. Krasovskii, *Some Problems in the Stability Theory of Motion*, Fiziko-Tekhnicheskikh, Moscow, Russia, 1959.
[13] A. S. Bychkov and D. Khusainov, "An exponential decay of solutions of neutral type stochastic equations," *Random Operators and Stochastic Equations*, vol. 3, no. 3, pp. 245–256, 1995.
[14] I. A. Dzhalladova, "Investigation of the stabilization of a mathematical model of a dynamical system with random action in the case of resonance," *Ukrainian Mathematical Journal*, vol. 49, no. 9, pp. 1324–1329, 1997, translated from Ukrans ki Matematichni Zhurnal, vol. 49, no. 9, pp. 1177–1181.
[15] I. A. Dzhalladova, "Investigation of the stabilization of a mathematical model of a dynamical system with random influence in the resonance case," *Ukrainian Mathematical Journal*, vol. 50, no. 8, pp. 1299–1307, 1998, translated from Ukrans ki Matematichni Zhurnal, vol. 50, no. 8, pp. 1137–1143.
[16] D. Ya. Khusainov and A. S. Bychkov, "Stability of stochastic systems of neutral type with small deviations of the argument," *Differential Equations*, vol. 28, no. 12, pp. 2060–2069, 1992.
[17] F. L. Lewis, *Optimal Estimation*, With an Introduction to Stochastic Control Theory, Wiley-Interscience Publication, John Wiley & Sons, New York, NY, USA, 1986.
[18] D. G. Korenevskiĭ, *Stability of Dynamical Systems under Random Perturbations of Parameters*, Algebraic Criteria, Naukova Dumka, Kiev, Ukraine, 1989.
[19] I. G. Nechaeva and D. Y. Khusainov, "Obtaining estimates for the stability of solutions of stochastic functional-differential equations," *Differential Equations*, vol. 28, no. 3, p. 405–414, 547, 1992.
[20] J.-H. Park and S. Won, "A note on stability of neutral delay-differential systems," *Journal of the Franklin Institute*, vol. 336, no. 3, pp. 543–548, 1999.
[21] V. Rezvan, *Absolute Stability Automatic System with Delay*, Nauka, Moscow, Russia, 1983.
[22] M. L. Sverdan, E. F. Tsarkov, and V. K. Yasinskii, *Stability in Stochastic Modeling of Complex Dynamical Systems*, Svyatyn, "Nad Prutom", 1996.
[23] E. F. Tsarkov, *Random Perturbations of Functional Difefrential Equations*, Zinatne, Riga, Latvia, 1989.

to Contents, page 1

[24] E. F. Tsarkov and V. K. Yasinskii, *Quasi-Linear Stochastic Functional Differential Equations*, Orientir, Riga, Latvia, 1992.

[25] X.-x. Liu and B. Xu, "A further note on stability criterion of linear neutral delay-differential systems," *Journal of the Franklin Institute*, vol. 343, no. 6, pp. 630–634, 2006.

[26] M. A. Ajzerman and F. R. Gantmacher, *Absolute Stability Regulations System*, Science Publishing House, Moscow, Russia, 1963.

[27] A. H. Gelig, G. A. Leonov, and V. A. Jakubovič, *The Stability of Nonlinear Systems with a Nonunique Equilibrium State*, Nauka, Moscow, Russia, 1978.

[28] A. K. Gelig and G. A. Leonov, *Nonlinear Systems*, Nauka, Moscow, Russia, 1978.

[29] A. D. Myshkis, *Linear Differential Equations with Delayed Argument*, Nauka, Moscow, Russia, 2nd edition, 1972.

[30] E. Boukas, *Deterministic and Stochastic Time Delay Systems*, Birkhäuser, Boston, Mass, USA, 2002.

[31] X. Liao and P. Yu, *Absolute Stability of Nonlinear Control Systems*, vol. 25 of *Mathematical Modelling: Theory and Applications*, Springer, New York, NY, USA, 2nd edition, 2008.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 754701, 16 pages
doi:10.1155/2011/754701

Research Article

Existence and Asymptotic Behavior of Positive Solutions of Functional Differential Equations of Delayed Type

J. Diblík and M. Kúdelčíková

Department of Mathematics, Faculty of Science, University of Žilina, 010 26 Žilina, Slovakia

Correspondence should be addressed to J. Diblík, josef.diblik@fpv.uniza.sk

Received 30 September 2010; Accepted 14 October 2010

Academic Editor: Elena Braverman

Solutions of the equation $\dot{y}(t) = -f(t, y_t)$ are considered for $t \to \infty$. The existence of two classes of positive solutions which are asymptotically different is proved using the retract method combined with Razumikhin's technique. With the aid of two auxiliary linear equations, which are constructed using upper and lower linear functional estimates of the right-hand side of the equation considered, inequalities for both types of positive solutions are given as well.

1. Introduction

Let $C([a,b], \mathbb{R}^n)$, where $a, b \in \mathbb{R}$, $a < b$, be the Banach space of the continuous mappings from the interval $[a,b]$ into $\mathbb{R}^n$ equipped with the supremum norm

$$\|\psi\|_C = \sup_{\theta \in [a,b]} \|\psi(\theta)\|, \quad \psi \in C([a,b], \mathbb{R}^n), \tag{1.1}$$

where $\|\cdot\|$ is the maximum norm in $\mathbb{R}^n$. In the case of $a = -r < 0$ and $b = 0$, we will denote this space as C_r^n, that is,

$$C_r^n := C([-r, 0], \mathbb{R}^n). \tag{1.2}$$

If $\sigma \in \mathbb{R}^n$, $A \geq 0$, and $y \in C([\sigma - r, \sigma + A], \mathbb{R}^n)$, then, for each $t \in [\sigma, \sigma + A]$, we define $y_t \in C_r^n$ by $y_t(\theta) = y(t + \theta)$, $\theta \in [-r, 0]$.

to Contents, page 1

The present article is devoted to the problem of the existence of two classes of asymptotically different positive solutions of the delayed equation

$$\dot{y}(t) = -f(t, y_t), \tag{1.3}$$

for $t \to +\infty$, where $f : \Omega \to \mathbb{R}$ is a continuous quasibounded functional that satisfies a local Lipschitz condition with respect to the second argument and Ω is an open subset in $\mathbb{R} \times C_r^1$ such that conditions which use f are well defined.

The main supposition of our investigation is that the right-hand side of (1.3) can be estimated as follows:

$$C_A(t) y_t(-r) \le f(t, y_t) \le C_B(t) y_t(-r), \tag{1.4}$$

where $(t, y_t) \in \Omega$, and $C_A, C_B : [t_0 - r, \infty) \to \mathbb{R}^+ := (0, \infty)$, $t_0 \in \mathbb{R}$ are continuous functions satisfying

$$0 < C_A(t) \le C_B(t) \le \frac{1}{(re)}, \quad t \in [t_0 - r, \infty), \tag{1.5}$$

$$\int_{t_0 - r}^{\infty} C_B(t) \mathrm{d}t < 1. \tag{1.6}$$

Quite lots of investigations are devoted to existence of positive solutions of different classes of equations (we mention at least monographs [1–6] and papers [7–12]). The investigation of two classes of asymptotically different solutions of (1.3) has been started in the paper [13] using a monotone iterative technique and a retract principle. Assumptions of results obtained are too cumbersome and are applied to narrow classes of equations. In the presented paper we derive more general statements under weaker conditions. This progress is related to more general inequalities (1.4) for the right-hand side of (1.3) which permit to omit utilization of properties of solutions of transcendental equations used in [13].

1.1. Ważewski's Principle

In this section we introduce Ważewski's principle for a system of retarded functional differential equations

$$\dot{y}(t) = F(t, y_t), \tag{1.7}$$

where $F : \Omega^* \mapsto \mathbb{R}^n$ is a continuous quasibounded map which satisfies a local Lipschitz condition with respect to the second argument and Ω^* is an open subset in $\mathbb{R} \times C_r^n$. We recall that the functional F is quasibounded if F is bounded on every set of the form $[t_1, t_2] \times C_{rL}^n \subset \Omega^*$, where $t_1 < t_2$, $C_{rL}^n := C([-r, 0], L)$ and L is a closed bounded subset of $\mathbb{R}^n$ (compare [2, page 305]).

In accordance with [14], a function $y(t)$ is said to be a *solution of system* (1.7) *on* $[\sigma - r, \sigma + A)$ if there are $\sigma \in \mathbb{R}$ and $A > 0$ such that $y \in C([\sigma - r, \sigma + A), \mathbb{R}^n)$, $(t, y_t) \in \Omega^*$, and $y(t)$ satisfies the system (1.7) for $t \in [\sigma, \sigma + A)$. For a given $\sigma \in \mathbb{R}$, $\varphi \in C$, we say $y(\sigma, \varphi)$ is a solution

to Contents, page 1

of the system (1.7) through $(\sigma, \varphi) \in \Omega^*$ if there is an $A > 0$ such that $y(\sigma, \varphi)$ is a solution of the system (1.7) on $[\sigma - r, \sigma + A)$ and $y_\sigma(\sigma, \varphi) = \varphi$. In view of the above conditions, each element $(\sigma, \varphi) \in \Omega^*$ determines a unique solution $y(\sigma, \varphi)$ of the system (1.7) through $(\sigma, \varphi) \in \Omega^*$ on its maximal interval of existence $I_{\sigma,\varphi} = [\sigma, a)$, $\sigma < a \leq \infty$ which depends continuously on initial data [14]. A solution $y(\sigma, \varphi)$ of the system (1.7) is said to be *positive* if

$$y_i(\sigma, \varphi) > 0 \tag{1.8}$$

on $[\sigma - r, \sigma] \cup I_{\sigma,\varphi}$ for each $i = 1, 2, \ldots, n$. A nontrivial solution $y(\sigma, \varphi)$ of the system (1.7) is said to be *oscillatory* on $I_{\sigma,\varphi}$ (under condition $I_{\sigma,\varphi} = [\sigma, \infty)$) if (1.8) does not hold on any subinterval $[\sigma_1, \infty) \subset [\sigma, \infty)$, $\sigma_1 \geq \sigma$.

As a method of proving the existence of positive solutions of (1.3), we use Ważewski's retract principle which was first introduced by Ważewski [15] for ordinary differential equations and later extended to retarded functional differential equations by Rybakowski [16] and which is widely applicable to concrete examples. A summary of this principle is given below.

As usual, if a set $\omega \subset \mathbb{R} \times \mathbb{R}^n$, then $\operatorname{int} \omega$ and $\partial\omega$ denote the interior and the boundary of ω, respectively.

Definition 1.1 (see [16]). Let the continuously differentiable functions $l_i(t, y)$, $i = 1, 2, \ldots, p$ and $m_j(t, y)$, $j = 1, 2, \ldots, q$, $p^2 + q^2 > 0$ be defined on some open set $\omega_0 \subset \mathbb{R} \times \mathbb{R}^n$. The set

$$\omega^* = \{(t, y) \in \omega_0 : l_i(t, y) < 0, m_j(t, y) < 0,\ i = 1, \ldots, p,\ j = 1, \ldots, q\} \tag{1.9}$$

is called a *regular polyfacial set* with respect to the system (1.7), provided that it is nonempty, if (α) to (γ) below hold.

(α) For $(t, \pi) \in \mathbb{R} \times C_r^n$ such that $(t + \theta, \pi(\theta)) \in \omega^*$ for $\theta \in [-r, 0)$, we have $(t, \pi) \in \Omega^*$.

(β) For all $i = 1, 2, \ldots, p$, all $(t, y) \in \partial\omega^*$ for which $l_i(t, y) = 0$, and all $\pi \in C_r^n$ for which $\pi(0) = y$ and $(t + \theta, \pi(\theta)) \in \omega^*$, $\theta \in [-r, 0)$. It follows that $Dl_i(t, y) > 0$, where

$$Dl_i(t, y) \equiv \sum_{k=1}^{n} \frac{\partial l_i(t, y)}{\partial y_k} f_k(t, \pi) + \frac{\partial l_i(t, y)}{\partial t}. \tag{1.10}$$

(γ) For all $j = 1, 2, \ldots, q$, all $(t, y) \in \partial\omega^*$ for which $m_j(t, y) = 0$, and all $\pi \in C_r^n$ for which $\pi(0) = y$ and $(t + \theta, \pi(\theta)) \in \omega^*$, $\theta \in [-r, 0)$. It follows that $Dm_j(t, y) < 0$, where

$$Dm_j(t, y) \equiv \sum_{k=1}^{n} \frac{\partial m_j(t, y)}{\partial y_k} f_k(t, \pi) + \frac{\partial m_j(t, y)}{\partial t}. \tag{1.11}$$

The elements $(t, \pi) \in \mathbb{R} \times C_r^n$ in the sequel are assumed to be such that $(t, \pi) \in \Omega^*$.

In the following definition, a set ω^* is an arbitrary set without any connection with a regular polyfacial set ω^* defined by (1.9) in Definition 1.1.

to Contents, page 1

Definition 1.2. A *system of initial functions* p_{A,ω^*} with respect to the nonempty sets A and ω^*, where $A \subset \overline{\omega^*} \subset \mathbb{R} \times \mathbb{R}^n$ is defined as a continuous mapping $p : A \longrightarrow C_r^n$ such that (α) and (β) below hold.

(α) If $z = (t, y) \in A \cap \operatorname{int} \omega^*$, then $(t + \theta, p(z)(\theta)) \in \omega^*$ for $\theta \in [-r, 0]$.

(β) If $z = (t, y) \in A \cap \partial\omega^*$, then $(t + \theta, p(z)(\theta)) \in \omega^*$ for $\theta \in [-r, 0)$ and $(t, p(z)(0)) = z$.

Definition 1.3 (see [17]). If $\mathcal{A} \subset \mathcal{B}$ are subsets of a topological space and $\pi : \mathcal{B} \longrightarrow \mathcal{A}$ is a continuous mapping from $\mathcal{B}$ onto $\mathcal{A}$ such that $\pi(p) = p$ for every $p \in \mathcal{A}$, then π is said to be a *retraction* of $\mathcal{B}$ onto $\mathcal{A}$. When a retraction of $\mathcal{B}$ onto $\mathcal{A}$ exists, $\mathcal{A}$ is called a *retract* of $\mathcal{B}$.

The following lemma describes the main result of the paper [16].

Lemma 1.4. *Let $\omega^* \subset \omega_0$ be a regular polyfacial set with respect to the system* (1.7), *and let W be defined as follows:*

$$W = \{(t, y) \in \partial\omega^* : m_j(t, y) < 0,\ j = 1, 2, \ldots, q\}. \tag{1.12}$$

Let $Z \subset W \cup \omega^*$ be a given set such that $Z \cap W$ is a retract of W but not a retract of Z. Then for each fixed system of initial functions p_{Z,ω^*}, there is a point $z_0 = (\sigma_0, y_0) \in Z \cap \omega^*$ such that for the corresponding solution $y(\sigma_0, p(z_0))(t)$ of (1.7), *one has*

$$(t, y(\sigma_0, p(z_0))(t)) \in \omega^* \tag{1.13}$$

for each $t \in D_{\sigma_0, p(z_0)}$.

Remark 1.5. When Lemma 1.4 is applied, a lot of technical details should be fulfilled. In order to simplify necessary verifications, it is useful, without loss of generality, to vary the first coordinate t in definition of the set ω^* in (1.9) within a half-open interval open at the right. Then the set ω^* is not open, but tracing the proof of Lemma 1.4, it is easy to see that for such sets it remains valid. Such possibility is used below. We will apply similar remark and explanation to sets of the type Ω, Ω^* which serve as domains of definitions of functionals on the right-hand sides of equations considered.

For continuous vector functions

$$\rho^* = (\rho_1^*, \rho_2^*, \ldots, \rho_n^*), \qquad \delta^* = (\delta_1^*, \delta_2^*, \ldots, \delta_n^*) : [t_0 - r, \infty) \longrightarrow \mathbb{R}^n, \tag{1.14}$$

with $\rho^*(t) \ll \delta^*(t)$ for $t \in [t_0 - r, \infty)$ (the symbol $\ll$ here and below means that $\rho_i^*(t) < \delta_i^*(t)$ for all $i = 1, 2, \ldots, n$), continuously differentiable on $[t_0, \infty)$, we define the set

$$\omega^* := \{(t, y) : t \in [t_0, \infty), \rho^*(t) \ll y \ll \delta^*(t)\}. \tag{1.15}$$

In the sequel, we employ the following result from [18, Theorem 1], which is proved with the aid of the retract technique combined with Razumikhin's approach.

to Contents, page 1

Theorem 1.6. *Let there be a* $p \in \{0, \ldots, n\}$ *such that*

(i) *if* $t \geq t_0$, $\phi \in C_r^n$ *and* $(t+\theta, \phi(\theta)) \in \omega^*$ *for any* $\theta \in [-r, 0)$, *then*

$$\begin{aligned} \left(\delta^{*i}\right)'(t) &< F^i(t,\phi), \quad \textit{when } \phi^i(0) = \delta^{*i}(t), \\ \left(\rho^{*i}\right)'(t) &> F^i(t,\phi), \quad \textit{when } \phi^i(0) = \rho^{*i}(t) \end{aligned} \tag{1.16}$$

for any $i = 1, 2, \ldots, p$, *(If* $p = 0$, *this condition is omitted.)*

(ii) *if* $t \geq t_0$, $\phi \in C_r^n$ *and* $(t+\theta, \phi(\theta)) \in \omega^*$ *for any* $\theta \in [-r, 0)$ *then*

$$\begin{aligned} \left(\rho^{*i}\right)'(t) &< F^i(t,\phi), \quad \textit{when } \phi^i(0) = \rho^{*i}(t), \\ \left(\delta^{*i}\right)'(t) &> F^i(t,\phi), \quad \textit{when } \phi^i(0) = \delta^{*i}(t) \end{aligned} \tag{1.17}$$

for any $i = p+1, p+2, \ldots, n$. *(If* $p = n$, *this condition is omitted.)*

Then, there exists an uncountable set $\mathcal{Y}$ *of solutions of* (1.7) *on* $[t_0 - r, \infty)$ *such that each* $y \in \mathcal{Y}$ *satisfies*

$$\rho^*(t) \ll y(t) \ll \delta^*(t), \quad t \in [t_0 - r, \infty). \tag{1.18}$$

1.2. Structure of Solutions of a Linear Equation

In this section we focus our attention to structure of solutions of scalar linear differential equation of the type (1.3) with variable bounded delay of the form

$$\dot{x}(t) = -c(t)x(t - \tau(t)) \tag{1.19}$$

with continuous functions $c : [t_0 - r, \infty) \to \mathbb{R}^+$ and $\tau : [t_0, \infty) \to (0, r]$.

In accordance with above definitions of positive or oscillatory solutions, we call a solution of (1.19) oscillatory if it has arbitrary large zeros, otherwise it is called nonoscillatory (positive or negative).

Let us mention properties of (1.19) which will be used later. Theorem 13 from [19] describes sufficient conditions for existence of positive solutions of (1.19) with nonzero limit.

Theorem 1.7 (see [19, Theorem 13]). *Linear equation* (1.19) *has a positive solution with nonzero limit if and only if*

$$\int^{\infty} c(t) \mathrm{d}t < \infty. \tag{1.20}$$

to Contents, page 1

Remark 1.8. Tracing the proof of Theorem 1.7, we conclude that a positive solution $x = x(t)$ of (1.19) with nonzero limit exists on $[t_0 - r, \infty)$ if

$$\int_{t_0-r}^{\infty} c(t)\mathrm{d}t < 1. \tag{1.21}$$

The following theorem is a union of parts of results from [20] related to the structure formulas for solutions of (1.19).

Theorem 1.9. *Suppose the existence of a positive solution of* (1.19) *on* $[t_0 - r, \infty)$*. Then there exist two positive solutions* x_d *and* x_s *of* (1.19) *on* $[t_0 - r, \infty)$ *satisfying the relation*

$$\lim_{t\to\infty} \frac{x_s(t)}{x_d(t)} = 0 \tag{1.22}$$

such that every solution $x = x(t)$ *of* (1.19) *on* $[t_0 - r, \infty)$ *can be represented by the formula*

$$x(t) = Kx_d(t) + O(x_s(t)), \tag{1.23}$$

where the constant K *depends on* x*.*

The symbol O, applied in (1.23) and below, is the Landau order symbol frequently used in asymptotic analysis.

Moreover, Theorem 9 in [20] gives a possibility to replace the pair of solutions $x_d(t)$ and $x_s(t)$ in (1.23) by another pairs of solutions $\tilde{x}_d(t)$ and $\tilde{x}_s(t)$ if

$$\lim_{t\to\infty} \frac{\tilde{x}_s(t)}{\tilde{x}_d(t)} = 0 \tag{1.24}$$

as given in the following theorem.

Theorem 1.10. *Let* $\tilde{x}_d(t)$ *and* $\tilde{x}_s(t)$ *be positive solutions of* (1.19) *on* $[t_0 - r, \infty)$ *such that* (1.24) *holds. Then every solution* $x = x(t)$ *of* (1.19) *on* $[t_0 - r, \infty)$ *can be represented by the formula*

$$x(t) = K^*\tilde{x}_d(t) + O(\tilde{x}_s(t)), \tag{1.25}$$

where the constant K^* *depends on* x*.*

The next definition is based on the properties of solutions x_d, $\tilde{x}_d$, x_s, and $\tilde{x}_s$ described in Theorems 1.9 and 1.10.

Definition 1.11 (see [20, Definition 2]). Suppose that the positive solutions x_d and x_s of (1.19) on $[t_0 - r, \infty)$ satisfy the relation (1.22). Then, we call the solution x_d a *dominant* solution and the solution x_s a *subdominant* solution.

Due to linearity of (1.19), there are infinitely many dominant and subdominant solutions. Obviously, another pair of a dominant and a subdominant solutions is the pair $\tilde{x}_d(t)$, $\tilde{x}_s(t)$ in Theorem 1.10.

to Contents, page 1

2. Main Results

Let us consider two auxiliary linear equations:

$$\dot{x}(t) = -C_A(t)x(t-r), \tag{2.1}$$

$$\dot{z}(t) = -C_B(t)z(t-r), \tag{2.2}$$

where $r \in \mathbb{R}^+$ and C_A, C_B are positive continuous functions on $[t_0 - r, \infty)$, $t_0 \in \mathbb{R}$. According to the Theorems 1.7 and 1.9, both (2.1) and (2.2) have two types of positive solutions (subdominant and dominant). Let us denote them $x_d(t)$, $x_s(t)$ for (2.1) and $z_d(t)$, $z_s(t)$ for (2.2), respectively, such that

$$\lim_{t\to\infty} \frac{x_s(t)}{x_d(t)} = 0, \qquad \lim_{t\to\infty} \frac{z_s(t)}{z_d(t)} = 0. \tag{2.3}$$

Without loss of generality, we can suppose that $x_s(t) < x_d(t)$ and $z_s(t) < z_d(t)$ on $[t_0 - r, \infty)$.

2.1. Auxiliary Linear Result

The next lemma states that if $z_d(t)$, $z_s(t)$ are dominant and subdominant solutions for (2.2), then there are dominant and subdominant solutions $x_d^*(t)$, $x_s^*(t)$ for (2.1) satisfying certain inequalities.

Lemma 2.1. *Let* (1.5) *be valid. Let* $z_d(t)$, $z_s(t)$ *be dominant and subdominant solutions for* (2.2). *Then there are positive solutions* $x_s^*(t)$, $x_d^*(t)$ *of* (2.1) *on* $[t_0 - r, \infty)$ *such that:*

(a) $x_s^*(t) < z_s(t)$, $t \in [t_0 - r, \infty)$,

(b) $z_d(t) < x_d^*(t)$, $t \in [t_0 - r, \infty)$,

(c) $x_d^*(t)$ *and* $x_s^*(t)$ *are dominant and subdominant solutions for* (2.1).

Proof. (a) To prove the part (a), we employ Theorem 1.6 with $p = n = 1$; that is, we apply the case (i). Consider (2.1), set $F(t,\phi) := -C_A(t)\phi(-r)$, $\rho^*(t) := 0$, $\delta^*(t) := z_s(t)$, and assume (see the case (i)):

$$0 < \phi(\theta) < z_s(t+\theta), \quad \theta \in [-r, 0), \ \phi(0) = z_s(t), \ t \geq t_0. \tag{2.4}$$

to Contents, page 1

Now we have to verify the inequalities (1.16), that is, in our case:

$$\begin{aligned} F(t,\phi) - (\delta^*)'(t) &= -C_A(t)\phi(-r) - (\delta^*)'(t) \\ &= -C_A(t)\phi(-r) - z_s'(t) \\ &= -C_A(t)\phi(-r) + C_B(t)z_s(t-r) \\ &\geq \text{(we use (1.5))} \\ &\geq -C_B(t)\phi(-r) + C_B(t)z_s(t-r) \\ &> C_B(t)[z_s(t-r) - z_s(t-r)] = 0 \end{aligned} \tag{2.5}$$

and $F(t,\phi) > (\delta^*)'(t)$ if $t \in [t_0,\infty)$. Further, we have

$$-F(t,\phi) + (\rho^*)'(t) = C_A(t)\phi(-r) + 0 = C_A(t)\phi(-r) > 0 \tag{2.6}$$

and $F(t,\phi) < (\rho^*)'(t)$ if $t \in [t_0,\infty)$. Since both inequalities are fulfilled and all assumptions of Theorem 1.6 are satisfied for the case in question, there exists a solution $x_s^*(t)$ of (2.1) on $[t_0 - r,\infty)$ such that $x_s^*(t) < z_s(t)$ for $t \in [t_0 - r,\infty)$.

(b) To prove the part (b), we consider a solution $x = x_d^*(t)$ of the following initial problem:

$$\dot{x}(t) = -C_A(t)x(t-r), \quad t \in [t_0 - r,\infty), \tag{2.7}$$

$$x(t) = z_d(t), \quad t \in [t_0 - r, t_0]. \tag{2.8}$$

Now, let us define a function

$$W(t,x) = z_d(t) - x(t), \quad t \in [t_0 - r,\infty). \tag{2.9}$$

We find the sign of the full derivative of W along the trajectories of (2.7) if $t \in [t_0, t_0 + r]$:

$$\begin{aligned} \left.\frac{\mathrm{d}W(t,x)}{\mathrm{d}t}\right|_{t\in[t_0,t_0+r]} &= -C_B(t)z_d(t-r) + C_A(t)x(t-r) \\ &= \text{(due to (2.8))} \\ &= -C_B(t)z_d(t-r) + C_A(t)z_d(t-r) \\ &\leq [C_A(t) - C_B(t)]z_d(t-r) \leq \text{(due to (1.5))} \leq 0. \end{aligned} \tag{2.10}$$

It means that function W is nonincreasing and it holds

$$\begin{aligned} W(t_0, x(t_0)) &= z_d(t_0) - x(t_0) = z_d(t_0) - z_d(t_0) \\ &= 0 \geq W(t_0+\varepsilon, x(t_0+\varepsilon)) = z_d(t_0+\varepsilon) - x(t_0+\varepsilon), \quad \varepsilon \in [0,r], \end{aligned} \tag{2.11}$$

to Contents, page 1

and hence $z_d(t_0+\varepsilon) \le x(t_0+\varepsilon)$. It will be showed that this inequality holds also for every $t > t_0 + r$.

On the contrary, let us suppose that the inequality is not true, that is, there exists a point $t = t^{**}$ such that $z_d(t^{**}) > x(t^{**})$. Then there exists a point $t^* \in [t_0, t^{**})$ such that $z_d(t^*) < x(t^*)$, otherwise $z_d(t) \equiv x(t)$ on $[t_0, t^{**}]$. Without loss of generality, we can suppose that $x(t) \equiv z_d(t)$ on $[t_0, t^{***}]$ with a $t^{***} \in [t_0, t^*)$ and $x(t) > z_d(t)$ on (t^{***}, t^*). Then, there exists a point $t^\diamond \in (t^{***}, t^*)$ such that $x(t^\diamond) = K z_d(t^\diamond)$ for a constant $K > 1$ and

$$K z_d(t) > x(t), \quad \text{for } t \in [t_0, t^\diamond). \tag{2.12}$$

Hence, for a function $W^*(t,x)$ defined as $W^*(t,x) := K z_d(t) - x(t)$, $t \in [t_0, t^\diamond]$, we get

$$\begin{aligned} \left.\frac{\mathrm{d}W^*(t,x)}{\mathrm{d}t}\right|_{t=t^\diamond} &= K(-C_B(t) z_d(t-r)) + C_A(t) x(t-r) \\ &< (\text{due to } (2.12)) \\ &< K(-C_B(t) z_d(t-r)) + C_A(t) K z_d(t-r) \\ &= K z_d(t-r)[C_A(t) - C_B(t)] \le (\text{by } (1.5)) \le 0. \end{aligned} \tag{2.13}$$

It means that $K z_d(t) < x(t)$ on a right-hand neighborhood of $t^\diamond$. This is a contradiction with inequality

$$z_d(t) < K z_d(t) < x(t), \tag{2.14}$$

hence it is proved that the existence of a solution $x_d^*(t)$ satisfies $z_d(t) < x_d^*(t)$ on $[t_0 - r, \infty)$.

(c) To prove the part (c), we consider $\lim_{t\to\infty} x_s^*(t)/x_d^*(t)$. Due to (a) and (b), we get

$$0 \le \lim_{t\to\infty} \frac{x_s^*(t)}{x_d^*(t)} \le \lim_{t\to\infty} \frac{z_s(t)}{z_d(t)} = 0, \tag{2.15}$$

and $x_d^*(t)$ and $x_s^*(t)$ are (by Definition 1.11) dominant and subdominant solutions for (2.1). □

2.2. Existence of Positive Solutions of (1.3)

The next theorems state that there exist two classes of positive solutions of (1.3) such that graphs of each solution of the first class are between graphs of dominant solutions of (2.1) and (2.2), and graphs of each solution of the second class are between graphs of subdominant solutions of (2.1) and (2.2), respectively. It means that we prove there are two classes of asymptotically different positive solutions of (1.3). Without loss of generality (see Remark 1.5), we put $\Omega := [t_0, \infty) \times C_r^1$. In the following, we will use our main supposition (1.4); that is, we assume that for $(t, \phi) \in \Omega$ inequalities,

$$C_A(t)\phi(-r) \le f(t,\phi) \le C_B(t)\phi(-r) \tag{2.16}$$

hold, where ϕ is supposed to be positive.

to Contents, page 1

Theorem 2.2. *Let $f : \Omega \to \mathbb{R}$ be a continuous quasibounded functional. Let inequality* (1.5) *be valid, and* (2.16) *holds for any $(t, \phi) \in \Omega$ with $\phi(\theta) > 0$, $\theta \in [-r, 0]$. Let $x(t)$ be a positive solution of* (2.1) *on $[t_0 - r, \infty)$, and let $z(t)$ be a positive solution of* (2.2) *on $[t_0 - r, \infty)$ such that $x(t) < z(t)$ on $[t_0 - r, \infty)$. Then there exists an uncountable set $\mathcal{Y}$ of positive solutions of* (1.3) *on $[t_0 - r, \infty)$ such that each solution $y \in \mathcal{Y}$ satisfies*

$$x(t) < y(t) < z(t) \tag{2.17}$$

for $t \in [t_0 - r, \infty)$.

Proof. To prove this theorem, we employ Theorem 1.6 with $p = n = 1$; that is, we apply the case (i). Set $F(t, y_t) := -f(t, y_t)$, $\rho^*(t) := x(t)$, $\delta^*(t) := z(t)$; hence, the set ω^* will be defined as

$$\omega^* := \{(t, y) : t \in [t_0 - r, \infty),\ x(t) < y(t) < z(t)\}. \tag{2.18}$$

Now, we have to verify the inequalities (1.16). In our case

$$\begin{aligned}
F(t, \phi) - (\delta^*)'(t) &= -f(t, \phi) - (\delta^*)'(t) \\
&= -f(t, \phi) - z'(t) \\
&= -f(t, \phi) + C_B(t) z(t - r) \\
&\geq (\text{we use } (2.16)) \\
&\geq -C_B(t)\phi(-r) + C_B(t) z(t - r) \\
&> (\text{we use } (2.18) : \phi(-r) < z(t - r)) \\
&> C_B(t)[z(t - r) - z(t - r)] = 0, \\
-F(t, \phi) + (\rho^*)'(t) &= f(t, \phi) + (\rho^*)'(t) \\
&= f(t, \phi) + x'(t) \\
&= f(t, \phi) - C_A(t) x(t - r) \\
&\geq (\text{we use } (2.16)) \\
&\geq C_A(t)\phi(-r) - C_A(t) x(t - r) \\
&> (\text{we use } (2.18) : \phi(-r) > x(t - r)) \\
&> C_A(t)[x(t - r) - x(t - r)] = 0.
\end{aligned} \tag{2.19}$$

Therefore,

$$\begin{aligned}
F(t, \phi) - (\delta^*)'(t) &> 0, \\
-F(t, \phi) + (\rho^*)'(t) &> 0.
\end{aligned} \tag{2.20}$$

to Contents, page 1

Both inequalities (1.16) are fulfilled, and all assumptions of Theorem 1.6 are satisfied for the case in question. There exists class of positive solutions $\mathcal{Y}$ of (1.3) on $[t_0 - r, \infty)$ that for each solution $y \in \mathcal{Y}$ from this class it is satisfied that $x(t) < y(t) < z(t)$ for $t \in [t_0 - r, \infty)$. □

Corollary 2.3. *Let, in accordance with Lemma 2.1, $x_s(t)$ be the subdominant solution of* (2.1)*, and let $z_s(t)$ be the subdominant solution of* (2.2)*, that is, $x_s(t) < z_s(t)$ on $[t_0 - r, \infty)$. Then, there exists an uncountable set $\mathcal{Y}_s$ of positive solutions of* (1.3) *on $[t_0 - r, \infty)$ such that each solution $y_s \in \mathcal{Y}_s$ satisfies*

$$x_s(t) < y_s(t) < z_s(t). \tag{2.21}$$

If inequality (1.6) holds, then dominant solutions $x_d(t)$ of (2.1) and $z_d(t)$ of (2.2) have finite positive limits

$$\begin{aligned} C_x &:= \lim_{t\to\infty} x_d(t), \quad C_x > 0, \\ C_z &:= \lim_{t\to\infty} z_d(t), \quad C_z > 0. \end{aligned} \tag{2.22}$$

This is a simple consequence of positivity of solutions $x_d(t)$, $z_d(t)$ and properties of dominant and subdominant solutions (see Theorem 1.7, Remark 1.8, Theorem 1.9, formulas (1.22)–(1.25) and (2.3)). Then, due to linearity of (2.1) and (2.2), it is clear that there are dominant solutions $x_d(t)$, $z_d(t)$ of both equations such that $z_d(t) < x_d(t)$ on $[t_0 - r, \infty)$. In the following lemma, we without loss of generality suppose that $x_d(t)$ and $z_d(t)$ are such solutions and their initial functions are nonincreasing on initial interval $[t_0 - r, t_0]$. We will need constants M and L satisfying

$$\begin{aligned} M > M^* &:= \frac{x_d(t_0 - r)}{C_z}, \\ L > L^* &:= \frac{M z_d(t_0 - r)}{C_x}. \end{aligned} \tag{2.23}$$

Lemma 2.4. *Let $f : \Omega \to \mathbb{R}$ be a continuous quasibounded functional. Let inequalities* (1.5) *and* (1.6) *be valid, and* (2.16) *holds for any $(t, \phi) \in \Omega$ with $\phi(\theta) > 0$, $\theta \in [-r, 0]$. Let $x_d(t)$, $t \in [t_0 - r, \infty)$ be a dominant solution of* (2.1)*, nonincreasing on $[t_0 - r, t_0]$, and let $z_d(t)$, $t \in [t_0 - r, \infty)$ be a dominant solution of* (2.2)*, nonincreasing on $[t_0 - r, t_0]$, such that $z_d(t) < x_d(t)$, $t \in [t_0 - r, \infty)$. Then there exists another dominant solution $z_d^*(t)$ of* (2.2) *and a positive solution $y = y_d(t)$ of* (1.3) *on $[t_0 - r, \infty)$ such that it holds that*

$$x_d(t) < y_d(t) < z_d^*(t) \tag{2.24}$$

for $t \in [t_0 - r, \infty)$ and $z_d^*(t) = M z_d(t)$.

to Contents, page 1

Proof. Both dominant solutions $x_d(t)$ and $z_d(t)$, of (2.1) and (2.2), respectively, have nonzero positive limits C_x and C_z. From linearity of (2.1) and (2.2), it follows that solutions multiplied by an arbitrary constant are also solutions of (2.1) and (2.2), respectively. It holds that

$$z_d^*(t_0 - r) = Mz_d(t_0 - r) \geq Mz_d(t) = z_d^*(t) > MC_z > x_d(t_0 - r) \geq x_d(t), \tag{2.25}$$

where $t \in [t_0 - r, \infty)$.

Now, we define the set ω^* in the same way as (2.18) in the proof of Theorem 2.2, but with $x_d(t)$ instead of $x(t)$ and with $z_d^*(t)$ instead of $z(t)$, that is,

$$\omega^* := \left\{ (t, y) : t \in [t_0 - r, \infty), x_d(t) < y(t) < z_d^*(t) \right\}. \tag{2.26}$$

According to the Theorem 2.2 (with $x_d(t)$ instead of $x(t)$ and with $z_d^*(t)$ instead of $z(t)$), it is visible that there exists a positive solution $y = y_d(t)$ of (1.3) satisfying

$$x_d(t) < y_d(t) < z_d^*(t), \tag{2.27}$$

where $t \in [t_0, \infty)$; that is, inequalities (2.24) hold. □

Theorem 2.5. *Let all suppositions of Lemma 2.4 be valid, and let $y_d(t)$ be a solution of (1.3) satisfying inequalities (2.24). Then, there exists a positive solution $x_d^{**}(t)$ of (2.1) on $[t_0 - r, \infty)$ satisfying*

$$z_d(t) < y_d(t) < x_d^{**}(t), \tag{2.28}$$

where $x_d^{**}(t) = Lx_d(t)$ and $t \in [t_0 - r, \infty)$.

Proof. Multiplying solution $x_d(t)$ by the constant L, we have

$$Lx_d(t) > LC_x > Mz_d(t_0 - r). \tag{2.29}$$

Using (2.29) and (2.24), we get

$$x_d^{**}(t) = Lx_d(t) > Mz_d(t_0 - r) = z_d^*(t_0 - r) > z_d^*(t) > y_d(t) > x_d(t) > z_d(t), \tag{2.30}$$

where $t \in [t_0 - r, \infty)$. Hence, there exists a solution $y_d(t)$ of (1.3) such that inequalities (2.28) hold. □

2.3. Asymptotically Different Behavior of Positive Solutions of (1.3)

Somewhat reformulating the statement of Theorem 2.5, we can define a class of positive solutions $\mathcal{Y}_d$ of (1.3) such that every solution $y_d \in \mathcal{Y}_d$ is defined on $[t_0 - r, \infty)$ and satisfies

$$Cz_d(t) < y_d(t) < Cx_d^{**}(t), \tag{2.31}$$

to Contents, page 1

where $t \in [t_0 - r, \infty)$ for a positive constant C and, for every positive constant C, there exists a solution $y_d \in \mathcal{Y}_d$ satisfying (2.31) on $[t_0 - r, \infty)$.

The following theorem states that positive solutions $y_s(t)$ and $y_d(t)$ of (1.3) have a different order of vanishing.

Theorem 2.6. *Let all the assumptions of Corollary 2.3 and Theorem 2.5 be met. Then there exist two classes $\mathcal{Y}_s$ and $\mathcal{Y}_d$ of positive solutions of (1.3) described by inequalities (2.21) and (2.31). Every two solutions y_s, y_d, such that $y_s \in \mathcal{Y}_s$ and $y_d \in \mathcal{Y}_d$, have asymptotically different behavior, that is,*

$$\lim_{t \to +\infty} \frac{y_s(t)}{y_d(t)} = 0. \tag{2.32}$$

Proof. Let the solution $y_s(t)$ be the one specified in Corollary 2.3 and the solution $y_d(t)$ specified by (2.31) with a positive constant C. Now let us verify that (2.32) holds. With the aid of inequalities (2.21) and (2.31), we get

$$0 \le \lim_{t \to +\infty} \frac{y_s(t)}{y_d(t)} \le \lim_{t \to +\infty} \frac{z_s(t)}{C z_d(t)} = 0 \tag{2.33}$$

in accordance with (1.22), since $z_s(t)$ and $z_d(t)$ are positive (subdominant and dominant) solutions of linear equation (2.2). □

Another final statement, being a consequence of Lemma 2.1 and Theorems 2.2 and 2.5, is the following.

Theorem 2.7. *Let $f : \Omega \to \mathbb{R}$ be a continuous quasibounded functional. Let inequalities (1.5) and (1.6) be valid, and (2.16) holds for any $(t, \phi) \in \Omega$ with $\phi(\theta) > 0$, $\theta \in [-r, 0]$. Then on $[t_0 - r, \infty)$ there exist*

(a) *dominant and subdominant solutions $x_d(t)$, $x_s(t)$ of (2.1),*

(b) *dominant and subdominant solutions $z_d(t)$, $z_s(t)$ of (2.2),*

(c) *solutions $y_d(t)$, $y_s(t)$ of (1.3)*

such that

$$0 < x_s(t) < y_s(t) < z_s(t) < z_d(t) < y_d(t) < x_d(t), \tag{2.34}$$

$$\lim_{t \to \infty} \frac{x_s(t)}{x_d(t)} = \lim_{t \to \infty} \frac{z_s(t)}{z_d(t)} = \lim_{t \to \infty} \frac{y_s(t)}{y_d(t)} = 0. \tag{2.35}$$

Example 2.8. Let (1.3) be reduced to

$$\dot{y}(t) = -f(t, y_t) := -3t \exp\left(-3t + \frac{1}{2}\cos(t y(t-1))\right) \cdot y(t-1), \tag{2.36}$$

to Contents, page 1

and let auxiliary linear equations (2.1) and (2.2) be reduced to

$$\dot{x}(t) = -4t\exp(2-4t)\cdot x(t-1), \tag{2.37}$$

$$\dot{z}(t) = -2t\exp(1-2t)\cdot z(t-1), \tag{2.38}$$

that is,

$$C_A(t) := 4t\exp(2-4t), \qquad C_B(t) := 2t\exp(1-2t), \qquad r = 1. \tag{2.39}$$

Let t_0 be sufficiently large. Inequalities (1.5), (1.6), and (2.16) hold. In view of linearity and by Remark 1.8, we conclude that there exist dominant solutions $x_d(t)$ of (2.37) and $z_d(t)$ of (2.38) such that

$$\lim_{t\to\infty} x_d(t) = 11, \quad \lim_{t\to\infty} z_d(t) = 2, \quad z_d(t) < x_d(t), \; t \in [t_0-1,\infty). \tag{2.40}$$

Moreover, there exist subdominant solutions $x_s(t)$ of (2.37) and $z_s(t)$ of (2.38) such that $x_s(t) < z_s(t)$, $t \in [t_0-1,\infty)$ which are defined as

$$x_s(t) := \exp\left(-2t^2\right), \qquad z_s(t) := \exp\left(-t^2\right). \tag{2.41}$$

By Theorem 2.7, we conclude that there exist solutions $y_s(t)$ and $y_d(t)$ of (2.36) satisfying inequalities (2.34), and (without loss of generality) inequalities

$$0 < x_s(t) = \exp\left(-2t^2\right) < y_s(t) < z_s(t) = \exp\left(-t^2\right) < 1 \le z_d(t) < y_d(t) < 10 \le x_d(t) \tag{2.42}$$

hold on $[t_0-1,\infty)$.

3. Conclusions and Open Problems

The following problems were not answered in the paper and present interesting topics for investigation.

Open Problem 3.1. In Lemma 2.4 and Theorems 2.5–2.7 we used the convergence assumption (1.6) being, without loss of generality, equivalent to

$$\int^{\infty} C_B(t)\mathrm{d}t < \infty. \tag{3.1}$$

It is an open question whether similar results could be proved if the integral is divergent, that is, if

$$\int^{\infty} C_B(t)\mathrm{d}t = \infty. \tag{3.2}$$

to Contents, page 1

Open Problem 3.2. Dominant and subdominant solutions are used for representation of family of all solutions of scalar linear differential delayed equation, for example, by formula (1.25). Investigation in this line of the role of solutions $y_d(t)$ and $y_s(t)$ of (1.3) (see Theorems 2.6 and 2.7) is an important question. Namely, it seems to be an interesting question to establish sufficient conditions for the right-hand side of (1.3) such that its every solution $y = y(t)$ can be represented on $[t_0 - r, \infty)$ by the formula

$$y(t) = Ky_d(t) + O\left(y_s(t)\right), \tag{3.3}$$

where the constant K depends only on $y(t)$.

Open Problem 3.3. The notions dominant and subdominant solutions are in the cited papers defined for scalar differential delayed equations only. It is a rather interesting question if the results presented can be enlarged to systems of differential delayed equations.

Remark 3.4. Except for papers and books mentioned in this paper we refer, for example, to sources [21–23], treating related problems as well. Note that the topic is connected with similar questions for discrete equations (e.g., [24–27]).

Acknowledgments

This research was supported by the Grant no. 1/0090/09 of the Grant Agency of Slovak Republic (VEGA) and by the Project APVV-0700-07 of Slovak Research and Development Agency.

References

[1] R. P. Agarwal, M. Bohner, and W.-T. Li, *Nonoscillation and Oscillation: Theory for Functional Differential Equations*, vol. 267 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 2004.

[2] R. D. Driver, *Ordinary and Delay Differential Equations*, Springer, New York, NY, USA, 1977.

[3] L. H. Erbe, Q. Kong, and B. G. Zhang, *Oscillation Theory for Functional-Differential Equations*, vol. 190 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 1995.

[4] K. Gopalsamy, *Stability and Oscillations in Delay Differential Equations of Population Dynamics*, vol. 74 of *Mathematics and Its Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1992.

[5] I. Győri and G. Ladas, *Oscillation Theory of Delay Differential Equations*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, NY, USA, 1991.

[6] V. Kolmanovskiĭ and A. Myshkis, *Applied Theory of Functional-Differential Equations*, vol. 85 of *Mathematics and Its Applications (Soviet Series)*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1992.

[7] B. Dorociaková and R. Olach, "Existence of positive solutions of delay differential equations," *Tatra Mountains Mathematical Publications*, vol. 43, pp. 63–70, 2009.

[8] E. Kozakiewicz, "Über das asymptotische Verhalten der nichtschwingenden Lösungen einer linearen Differentialgleichung mit nacheilendem Argument," *Wissenschaftliche Zeitschrift Der Humboldt Universitat Zu Berlin, Mathematisch-Naturwissenschaftliche Reihe*, vol. 13, no. 4, pp. 577–589, 1964.

[9] E. Kozakiewicz, "Zur Abschätzung des Abklingens der nichtschwingenden Lösungen einer linearen Differentialgleichung mit nacheilendem Argument," *Wissenschaftliche Zeitschrift Der Humboldt Universitat Zu Berlin, Mathematisch-Naturwissenschaftliche Reihe*, vol. 15, pp. 675–676, 1966.

[10] E. Kozakiewicz, "Über die nichtschwingenden Lösungen einer linearen Differentialgleichung mit nacheilendem Argument," *Mathematische Nachrichten*, vol. 32, pp. 107–113, 1966.

to Contents, page 1

[11] M. Pituk, "Special solutions of functional differential equations," *Studies of the University of Žilina. Mathematical Series*, vol. 17, no. 1, pp. 115–122, 2003.
[12] Yu. A. Ryabov, "Certain asymptotic properties of linear systems with small time lag," *Trudy Sem. Teor. Diff. Uravnenii s Otklon. Argumentom Univ. Druzby Narodov Patrisa Lumumby*, vol. 3, pp. 153–165, 1965 (Russian).
[13] J. Diblík and M. Kúdelčíková, "Two classes of asymptotically different positive solutions of the equation $\dot{y}(t) = -f(t, y_t)$," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 70, no. 10, pp. 3702–3714, 2009.
[14] J. K. Hale and S. M. Verduyn Lunel, *Introduction to Functional-Differential Equations*, vol. 99 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 1993.
[15] T. Ważewski, "Sur un principe topologique de l'examen de l'allure asymptotique des intégrales des équations différentielles ordinaires," *Annales De La Societe Polonaise De Mathematique*, vol. 20, pp. 279–313, 19487.
[16] K. P. Rybakowski, "Ważewski's principle for retarded functional differential equations," *Journal of Differential Equations*, vol. 36, no. 1, pp. 117–138, 1980.
[17] V. Lakshmikantham and S. Leela, *Differential and Integral Inequalities—vol. I. Ordinary Differential Equations*, Academic Press, New York, NY, USA, 1969.
[18] J. Diblík, "A criterion for existence of positive solutions of systems of retarded functional-differential equations," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 38, no. 3, pp. 327–339, 1999.
[19] J. Diblík, "Behaviour of solutions of linear differential equations with delay," *Archivum Mathematicum*, vol. 34, no. 1, pp. 31–47, 1998.
[20] J. Diblík and M. Růžičková, "Asymptotic behavior of solutions and positive solutions of differential delayed equations," *Functional Differential Equations*, vol. 14, no. 1, pp. 83–105, 2007.
[21] J. Čermák, "On a linear differential equation with a proportional delay," *Mathematische Nachrichten*, vol. 280, no. 5-6, pp. 495–504, 2007.
[22] I. Culáková, L. Hanuštiaková, and R. Olach, "Existence for positive solutions of second-order neutral nonlinear differential equations," *Applied Mathematics Letters*, vol. 22, no. 7, pp. 1007–1010, 2009.
[23] L. K. Kikina and I. P. Stavroulakis, "A survey on the oscillation of solutions of first order delay difference equations," *Cubo*, vol. 7, no. 2, pp. 223–236, 2005.
[24] J. Baštinec and J. Diblík, "Subdominant positive solutions of the discrete equation $\Delta u(k + n) = -p(k)u(k)$," *Abstract and Applied Analysis*, vol. 2004, pp. 461–470, 2004.
[25] L. Berezansky, E. Braverman, and O. Kravets, "Nonoscillation of linear delay difference equations with positive and negative coefficients," *Journal of Difference Equations and Applications*, vol. 14, no. 5, pp. 495–511, 2008.
[26] I. Győri and L. Horváth, "Asymptotic constancy in linear difference equations: limit formulae and sharp conditions," *Advances in Difference Equations*, vol. 2010, Article ID 789302, 20 pages, 2010.
[27] S. Stević and K. S. Berenhaut, "The behavior of positive solutions of a nonlinear second-order difference equation," *Abstract and Applied Analysis*, vol. 2008, Article ID 653243, 8 pages, 2008.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 792689, 13 pages
doi:10.1155/2011/792689

Research Article

Existence Conditions for Bounded Solutions of Weakly Perturbed Linear Impulsive Systems

Alexander Boichuk,[1] Martina Langerová,[2] and Jaroslava Škoríková[2]

[1] *Intitute of Mathematics, National Academy of Science of Ukraine, 01601 Kyiv, Ukraine*
[2] *Department of Mathematics, University of Žilina, 01026 Žilina, Slovakia*

Correspondence should be addressed to Martina Langerová, martina.langerova@fpv.uniza.sk

Received 31 December 2010; Revised 1 July 2011; Accepted 1 July 2011

Academic Editor: Josef Diblík

The weakly perturbed linear nonhomogeneous impulsive systems in the form $\dot{x} = A(t)x + \varepsilon A_1(t)x + f(t)$, $t \in \mathbb{R}$, $t \notin \mathcal{T} := \{\tau_i\}_{\mathbb{Z}}$, $\Delta x|_{t=\tau_i} = \gamma_i + \varepsilon A_{1i}x(\tau_i-)$, $\tau_i \in \mathcal{T} \subset \mathbb{R}$, $\gamma_i \in \mathbb{R}^n$, and $i \in \mathbb{Z}$ are considered. Under the assumption that the generating system (for $\varepsilon = 0$) does not have solutions bounded on the entire real axis for some nonhomogeneities and using the Vishik-Lyusternik method, we establish conditions for the existence of solutions of these systems bounded on the entire real axis in the form of a Laurent series in powers of small parameter ε with finitely many terms with negative powers of ε, and we suggest an algorithm of construction of these solutions.

1. Introduction

In this contribution we study the problem of existence and construction of solutions of weakly perturbed linear differential systems with impulsive action bounded on the entire real axis. The application of the theory of differential systems with impulsive action (developed in [1–3]), the well-known results on the splitting index by Sacker [4] and by Palmer [5] on the Fredholm property of bounded solutions of linear systems of ordinary differential equations [6–9], the theory of pseudoinverse matrices [10] and results obtained in analyzing boundary-value problems for ordinary differential equations (see [10–12]), enables us to obtain existence conditions and to propose an algorithm for the construction of solutions bounded on the entire real axis of weakly perturbed linear impulsive differential systems.

to Contents, page 1

2. Initial Problem

We consider the problem of existence and construction of solutions bounded on the entire real axis of linear systems of ordinary differential equations with impulsive action at fixed points of time

$$\begin{aligned} \dot{x} &= A(t)x + f(t), \quad t \in \mathbb{R} \setminus \mathcal{T}, \\ \Delta x|_{t=\tau_i} &= \gamma_i, \quad \tau_i \in \mathcal{T}, \; i \in \mathbb{Z}, \end{aligned} \tag{2.1}$$

where $A \in BC_{\mathcal{T}}(\mathbb{R})$ is an $n \times n$ matrix of functions, $f \in BC_{\mathcal{T}}(\mathbb{R})$ is an $n \times 1$ vector function, $BC_{\mathcal{T}}(\mathbb{R})$ is the Banach space of real vector functions bounded on $\mathbb{R}$ and left-continuous for $t \in \mathbb{R}$ with discontinuities of the first kind at $t \in \mathcal{T} := \{\tau_i\}_{\mathbb{Z}}$ with the norm: $\|x\|_{BC_{\mathcal{T}}(\mathbb{R})} := \sup_{t\in\mathbb{R}}\|x(t)\|$, γ_i are n-dimensional column constant vectors: $\gamma_i \in \mathbb{R}^n$; $\cdots < \tau_{-2} < \tau_{-1} < \tau_0 = 0 < \tau_1 < \tau_2 < \cdots$, and $\Delta x|_{t=\tau_i} := x(\tau_i+) - x(\tau_i-)$.

The solution $x(t)$ of the system (2.1) is sought in the Banach space of n-dimensional bounded on $\mathbb{R}$ and piecewise continuously differentiable vector functions with discontinuities of the first kind at $t \in \mathcal{T}: \; x \in BC^1_{\mathcal{T}}(\mathbb{R})$.

Parallel with the nonhomogeneous impulsive system (2.1), we consider the corresponding homogeneous system

$$\dot{x} = A(t)x, \qquad \Delta x|_{t=\tau_i} = 0, \tag{2.2}$$

which is the homogeneous system without impulses, and let $X(t)$ be the fundamental matrix of (2.2) such that $X(0) = I$.

Assume that the homogeneous system (2.2) is exponentially dichotomous (e-dichotomous) [5, 10] on semiaxes $\mathbb{R}_- = (-\infty, 0]$ and $\mathbb{R}_+ = [0, \infty)$, that is, there exist projectors P and Q ($P^2 = P$, $Q^2 = Q$) and constants $K_i \geq 1$, $\alpha_i > 0$ ($i = 1, 2$) such that the following inequalities are satisfied:

$$\begin{aligned} \left\| X(t)PX^{-1}(s) \right\| &\leq K_1 e^{-\alpha_1(t-s)}, \quad t \geq s, \\ \left\| X(t)(I-P)X^{-1}(s) \right\| &\leq K_1 e^{-\alpha_1(s-t)}, \quad s \geq t, \; t, s \in \mathbb{R}_+, \\ \left\| X(t)QX^{-1}(s) \right\| &\leq K_2 e^{-\alpha_2(t-s)}, \quad t \geq s, \\ \left\| X(t)(I-Q)X^{-1}(s) \right\| &\leq K_2 e^{-\alpha_2(s-t)}, \quad s \geq t, \; t, s \in \mathbb{R}_-. \end{aligned} \tag{2.3}$$

to Contents, page 1

For getting the solution $x \in BC^1_{\mathcal{T}}(\mathbb{R})$ bounded on the entire axis, we assume that $t = 0 \notin \mathcal{T}$, that is, $x(0+) - x(0-) = \gamma_0 = 0$.

We use the following notation: $D = P - (I - Q)$; D^+ is a Moore-Penrose pseudoinverse matrix to D; P_D and P_{D^*} are $n \times n$ matrices (orthoprojectors) projecting $\mathbb{R}^n$ onto $N(D) = \ker D$ and onto $N(D^*) = \ker D^*$, respectively, that is, $P_D : \mathbb{R}^n \to N(D)$, $P_D^2 = P_D = P_D^*$, and $P_{D^*} : \mathbb{R}^n \to N(D^*)$, $P_{D^*}^2 = P_{D^*} = P_{D^*}^*$; $H(t) = [P_{D^*}Q]X^{-1}(t)$; $d = \operatorname{rank}[P_{D^*}Q] = \operatorname{rank}[P_{D^*}(I-P)]$ and $r = \operatorname{rank}[PP_D] = \operatorname{rank}[(I-Q)P_D]$.

The existence conditions and the structure of solutions of system (2.1) bounded on the entire real axis was analyzed in [13]. Here the following theorem was formulated and proved.

Theorem 2.1. *Assume that the linear nonhomogeneous impulsive differential system* (2.1) *has the corresponding homogeneous system* (2.2) *e-dichotomous on the semiaxes* $\mathbb{R}_- = (-\infty, 0]$ *and* $\mathbb{R}_+ = [0, \infty)$ *with projectors* P *and* Q*, respectively. Then the homogeneous system* (2.2) *has exactly* r *linearly independent solutions bounded on the entire real axis. If nonhomogeneities* $f \in BC_{\mathcal{T}}(\mathbb{R})$ *and* $\gamma_i \in \mathbb{R}^n$ *satisfy* d *linearly independent conditions*

$$\int_{-\infty}^{\infty} H_d(t) f(t) dt + \sum_{i=-\infty}^{\infty} H_d(\tau_i)\gamma_i = 0, \tag{2.4}$$

then the nonhomogeneous system (2.1) *possesses an* r*-parameter family of linearly independent solutions bounded on* $\mathbb{R}$ *in the form*

$$x(t, c_r) = X_r(t)c_r + \left(G\begin{bmatrix} f \\ \gamma_i \end{bmatrix}\right)(t), \quad \forall c_r \in \mathbb{R}^r. \tag{2.5}$$

Here, $H_d(t) = [P_{D^*}Q]_d X^{-1}(t)$ is a $d \times n$ matrix formed by a complete system of d linearly independent rows of matrix $H(t)$,

$$X_r(t) := X(t)[PP_D]_r = X(t)[(I-Q)P_D]_r \tag{2.6}$$

is an $n \times r$ matrix formed by a complete system of r linearly independent solutions bounded on $\mathbb{R}$ of homogeneous system (2.2), and $\left(G\left[\begin{smallmatrix} f \\ \gamma_i \end{smallmatrix}\right]\right)(t)$ is the generalized Green operator of

to Contents, page 1

the problem of finding bounded solutions of the nonhomogeneous impulsive system (2.1), acting upon $f \in BC_{\mathcal{T}}(\mathbb{R})$ and $\gamma_i \in \mathbb{R}^n$, defined by the formula

$$\left(G\begin{bmatrix} f \\ \gamma_i \end{bmatrix}\right)(t) = X(t)\begin{cases} \displaystyle\int_0^t PX^{-1}(s)f(s)ds - \int_t^{\infty}(I-P)X^{-1}(s)f(s)ds \\ \displaystyle\quad + \sum_{i=1}^{j} PX^{-1}(\tau_i)\gamma i - \sum_{i=j+1}^{\infty}(I-P)X^{-1}(\tau_i)\gamma_i \\ \displaystyle\quad + PD^{+}\left\{\int_{-\infty}^{0} QX^{-1}(s)f(s)ds + \int_0^{\infty}(I-P)X^{-1}(s)f(s)ds\right. \\ \displaystyle\qquad\left. + \sum_{i=-\infty}^{-1} QX^{-1}(\tau_i)\gamma_i + \sum_{i=1}^{\infty}(I-P)X^{-1}(\tau_i)\gamma_i\right\}, & t \geq 0; \\ \displaystyle\int_{-\infty}^t QX^{-1}(s)f(s)ds - \int_t^{0}(I-Q)X^{-1}(s)f(s)ds \\ \displaystyle\quad + \sum_{i=-\infty}^{-(j+1)} QX^{-1}(\tau_i)\gamma_i - \sum_{i=-j}^{-1}(I-Q)X^{-1}(\tau_i)\gamma_i \\ \displaystyle\quad + (I-Q)D^{+}\left\{\int_{-\infty}^{0} QX^{-1}(s)f(s)ds + \int_0^{\infty}(I-P)X^{-1}(s)f(s)ds\right. \\ \displaystyle\qquad\left. + \sum_{i=-\infty}^{-1} QX^{-1}(\tau_i)\gamma_i + \sum_{i=1}^{\infty}(I-P)X^{-1}(\tau_i)\gamma_i\right\}, & t \leq 0, \end{cases} \tag{2.7}$$

with the following property

$$\left(G\begin{bmatrix} f \\ \gamma_i \end{bmatrix}\right)(0-) - \left(G\begin{bmatrix} f \\ \gamma_i \end{bmatrix}\right)(0+) = \int_{-\infty}^{\infty} H(t)f(t)dt + \sum_{i=-\infty}^{\infty} H(\tau_i)\gamma_i. \tag{2.8}$$

These results are required to establish new conditions for the existence of solutions of weakly perturbed linear impulsive systems bounded on the entire real axis.

3. Perturbed Problems

Consider a weakly perturbed nonhomogeneous linear impulsive system in the form

$$\begin{gathered} \dot{x} = A(t)x + \varepsilon A_1(t)x + f(t), \quad t \in \mathbb{R} \setminus \mathcal{T}, \\ \Delta x|_{t=\tau_i} = \gamma_i + \varepsilon A_{1i}x(\tau_i-), \quad \tau_i \in \mathcal{T},\ \gamma_i \in \mathbb{R}^n,\ i \in \mathbb{Z}, \end{gathered} \tag{3.1}$$

where $A_1 \in BC_{\mathcal{T}}(\mathbb{R})$ is an $n \times n$ matrix of functions, A_{1i} are $n \times n$ constant matrices.

Assume that the condition of solvability (2.4) of the generating system (2.1) (obtained from system (3.1) for $\varepsilon = 0$) is not satisfied for all nonhomogeneities $f \in BC_{\mathcal{T}}(\mathbb{R})$ and $\gamma_i \in \mathbb{R}^n$, that is, system (2.1) does not have solutions bounded on the entire real axis. Therefore, we analyze whether the system (2.1) can be made solvable by introducing linear perturbations

to Contents, page 1

to the differential system and to the pulsed conditions. Also it is important to determine perturbations $A_1(t)$ and A_{1i} required to make the problem (3.1) solvable in the space of functions bounded on the entire real axis, that is, it is necessary to specify pertubations for which the corresponding homogeneous system

$$\begin{gathered} \dot{x} = A(t)x + \varepsilon A_1(t)x, \quad t \in \mathbb{R} \setminus \mathcal{T}, \\ \Delta x|_{t=\tau_i} = \varepsilon A_{1i} x(\tau_i -), \quad \tau_i \in \mathcal{T}, \ i \in \mathbb{Z}, \end{gathered} \tag{3.2}$$

turns into a system e-trichotomous or e-dichotomous on the entire real axis [10].

We show that this problem can be solved using the $d \times r$ matrix

$$B_0 = \int_{-\infty}^{\infty} H_d(t) A_1(t) X_r(t) dt + \sum_{i=-\infty}^{\infty} H_d(\tau_i) A_{1i} X_r(\tau_i -), \tag{3.3}$$

constructed with the coefficients of the system (3.1). The Vishik-Lyusternik method developed in [14] enables us to establish conditions under which a solution of impulsive system (3.1) can be represented by a function bounded on the entire real axis in the form of a Laurent series in powers of the small fixed parameter ε with finitely many terms with negative powers of ε.

We use the following notation: B_0^+ is the unique matrix pseudoinverse to B_0 in the Moore-Penrose sense, P_{B_0} is the $r \times r$ matrix (orthoprojector) projecting the space R^r to the null space $N(B_0)$ of the $d \times r$ matrix B_0, that is, $P_{B_0}: R^r \to N(B_0)$, and $P_{B_0^*}$ is the $d \times d$ matrix (orthoprojector) projecting the space $\mathbb{R}^d$ to the null space $N(B_0^*)$ of the $r \times d$ matrix B_0^* ($B_0^* = B^T$), that is, $P_{B_0^*} : \mathbb{R}^d \to N(B_0^*)$.

Now we formulate and prove a theorem that enables us to solve indicated problem.

Theorem 3.1. *Suppose that the system* (3.1) *satisfies the conditions imposed above, and the homogeneous system* (2.2) *is e-dichotomous on* $\mathbb{R}_+$ *and* $\mathbb{R}_-$ *with projectors* P *and* Q*, respectively. Let nonhomogeneities* $f \in BC_{\mathcal{T}}(\mathbb{R})$ *and* $\gamma_i \in \mathbb{R}^n$ *be given such that the condition* (2.4) *is not satisfied and the generating system* (2.1) *does not have solutions bounded on the entire real axis. If*

$$P_{B_0^*} = 0, \tag{3.4}$$

then the system (3.2) *is e-trichotomous on* $\mathbb{R}$ *and, for all nonhomogeneities* $f \in BC_{\mathcal{T}}(\mathbb{R})$ *and* $\gamma_i \in \mathbb{R}^n$*, the system* (3.1) *possesses at least one solution bounded on* $\mathbb{R}$ *in the form of a series*

$$x(t, \varepsilon) = \sum_{k=-1}^{\infty} \varepsilon^k x_k(t), \tag{3.5}$$

uniformly convergent for sufficiently small fixed $\varepsilon \in (0, \varepsilon_*]$.

to Contents, page 1

Here, ε_* is a proper constant characterizing the range of convergence of the series (3.5) and the coefficients $x_k(t)$ of the series (3.5) are determined from the corresponding impulsive systems as

$$x_k(t) = x_k(t, c_k) = X_r(t)c_k + \left(G\begin{bmatrix} A_1(\cdot)x_{k-1}(\cdot, c_{k-1}) \\ A_{1i}x(\tau_i-, c_{k-1}) \end{bmatrix} \right)(t) \quad \text{for } k = 1, 2, \dots,$$

$$c_k = -B_0^+ \left[\int_{-\infty}^{\infty} H_d(t)A_1(t)\left(G\begin{bmatrix} A_1(\cdot)x_{k-1}(\cdot, c_{k-1}) \\ A_{1i}x_{k-1}(\tau_i-, c_{k-1}) \end{bmatrix} \right)(t)dt \right.$$
$$\left. + \sum_{i=-\infty}^{\infty} H_d(\tau_i)A_{1i}\left(G\begin{bmatrix} A_1(\cdot)x_{k-1}(\cdot, c_{k-1}) \\ A_{1i}x_{k-1}(\cdot, c_{k-1}) \end{bmatrix} \right)(\tau_i-) \right],$$

$$x_{-1}(t) = x_{-1}(t, c_{-1}) = X_r(t)c_{-1}, \quad c_{-1} = B_0^+ \left\{ \int_{-\infty}^{\infty} H_d(t)f(t)dt + \sum_{i=-\infty}^{\infty} H_d(\tau_i-)\gamma_i \right\},$$

$$x_0(t) = x_0(t, c_0) = X_r(t)c_0 + \left(G\begin{bmatrix} A_1(\cdot)X_r(t)c_{-1} + f(\cdot) \\ \gamma_i + A_{1i}X_r(\tau_i-)c_{-1} \end{bmatrix} \right)(t),$$

$$c_0 = -B_0^+ \left[\int_{-\infty}^{\infty} H_d(t)A_1(t)\left(G\begin{bmatrix} A_1(\cdot)x_{-1}(\cdot, c_{-1}) + f(\cdot) \\ A_{1i}x_{-1}(\tau_i-, c_{-1}) + \gamma_i \end{bmatrix} \right)(t)dt \right.$$
$$\left. + \sum_{i=-\infty}^{\infty} H_d(\tau_i)A_{1i}\left(G\begin{bmatrix} A_1(\cdot)x_{-1}(\cdot, c_{-1}) + f(\cdot) \\ A_{1i}x_{-1}(\cdot, c_{-1}) + \gamma_i \end{bmatrix} \right)(\tau_i-) \right]. \tag{3.6}$$

Proof. We suppose that the problem (3.1) has a solution in the form of a Laurent series (3.5). We substitute this solution into the system (3.1) and equate the coefficients at the same powers of ε. The problem of determination of the coefficient $x_{-1}(t)$ of the term with ε^{-1} in series (3.5) is reduced to the problem of finding solutions of homogeneous system without impulses

$$\begin{aligned} \dot{x}_{-1} &= A(t)x_{-1}, \quad t \notin \mathcal{T}, \\ \Delta x_{-1}|_{t=\tau_i} &= 0, \quad i \in \mathbb{Z}, \end{aligned} \tag{3.7}$$

bounded on the entire real axis. According to the Theorem 2.1, the homogeneous system (3.7) possesses r-parameter family of solutions

$$x_{-1}(t, c_{-1}) = X_r(t)c_{-1} \tag{3.8}$$

bounded on the entire real axis, where c_{-1} is an r-dimensional vector column $c_{-1} \in \mathbb{R}^r$ and is determined from the condition of solvability of the problem used for determining the coefficient x_0 of the series (3.5).

to Contents, page 1

For ε^0, the problem of determination of the coefficient $x_0(t)$ of series (3.5) reduces to the problem of finding solutions of the following nonhomogeneous system:

$$\begin{gathered}\dot{x}_0 = A(t)x_0 + A_1(t)x_{-1} + f(t), \quad t \notin \mathcal{T}, \\ \Delta x_0|_{t=\tau_i} = A_{1i}x_{-1}(\tau_i-) + \gamma_i, \quad i \in \mathbb{Z},\end{gathered} \tag{3.9}$$

bounded on the entire real axis. According to the Theorem 2.1, the condition of solvability of this problem takes the form

$$\int_{-\infty}^{\infty} H_d(t)\left[A_1(t)X_r(t)c_{-1} + f(t)\right]dt + \sum_{i=-\infty}^{\infty} H_d(\tau_i)\left[A_{1i}X_r(\tau_i-)c_{-1} + \gamma_i\right] = 0. \tag{3.10}$$

Using the matrix B_0, we get the following algebraic system for $c_{-1} \in \mathbb{R}^r$:

$$B_0c_{-1} = -\int_{-\infty}^{\infty} H_d(t)f(t)dt + \sum_{i=-\infty}^{\infty} H_d(\tau_i-)\gamma_i, \tag{3.11}$$

which is solvable if and only if the condition

$$P_{B_0^*}\left\{\int_{-\infty}^{\infty} H_d(t)f(t)dt + \sum_{i=-\infty}^{\infty} H_d(\tau_i-)\gamma_i\right\} = 0 \tag{3.12}$$

is satisfied, that is, if

$$P_{B_0^*} = 0. \tag{3.13}$$

In this case, this algebraic system is solvable with respect to $c_{-1} \in \mathbb{R}^r$ within an arbitrary vector constant $P_{B_0}c(\forall c \in \mathbb{R}^r)$ from the null space of the matrix B_0, and one of its solutions has the form

$$c_{-1} = B_0^+\left\{\int_{-\infty}^{\infty} H_d(t)f(t)dt + \sum_{i=-\infty}^{\infty} H_d(\tau_i-)\gamma_i\right\}. \tag{3.14}$$

Therefore, under condition (3.4), the nonhomogeneous system (3.9) possesses an r-parameter set of solution bounded on $\mathbb{R}$ in the form

$$x_0(t, c_0) = X_r(t)c_0 + \left(G\begin{bmatrix} A_1(\cdot)x_{-1}(\cdot, c_{-1}) + f(\cdot) \\ \gamma_i + A_{1i}x_{-1}(\tau_i-, c_{-1}) \end{bmatrix}\right)(t), \tag{3.15}$$

where $(G[{}^*_*])(t)$ is the generalized Green operator (2.7) of the problem of finding bounded solutions of system (3.9), and c_0 is an r-dimensional constant vector determined in the next step of the process from the condition of solvability of the impulsive problem for coefficient $x_1(t)$.

to Contents, page 1

We continue this process by problem of determination of the coefficient $x_1(t)$ of the term with ε^1 in the series (3.5). It reduces to the problem of finding solutions of the system

$$\begin{aligned} \dot{x}_1 &= A(t)x_1 + A_1(t)x_0, \quad t \notin \mathcal{T}, \\ \Delta x_1|_{t=\tau_i} &= A_{1i}x_0(\tau_i-), \quad i \in \mathbb{Z}, \end{aligned} \tag{3.16}$$

bounded on the entire real axis. If the condition (3.4) is satisfied and by using the condition of solvability of this problem, that is,

$$\begin{aligned} &\int_{-\infty}^{\infty} H_d(t)A_1(t)\left[X_r(t)c_0 + \left(G\begin{bmatrix} A_1(\cdot)x_{-1}(\cdot,c_{-1}) + f(\cdot) \\ A_{1i}x_{-1}(\tau_i-,c_{-1}) + \gamma_i \end{bmatrix}\right)(t)\right]dt \\ &\quad + \sum_{i=-\infty}^{\infty} H_d(\tau_{i-})A_{1i}\left[X_r(\tau_i-)c_0 + \left(G\begin{bmatrix} A_1(\cdot)x_{-1}(\cdot,c_{-1}) + f(\cdot) \\ A_{1i}x_{-1}(\cdot,c_{-1}) + \gamma_i \end{bmatrix}\right)(\tau_i-)\right] = 0, \end{aligned} \tag{3.17}$$

we determine the vector $c_0 \in \mathbb{R}^r$ (within an arbitrary vector constant $P_{B_0}c, \forall c \in \mathbb{R}^r$) as

$$\begin{aligned} c_0 = -B_0^+ &\left[\int_{-\infty}^{\infty} H_d(t)A_1(t)\left(G\begin{bmatrix} A_1(\cdot)x_{-1}(\cdot,c_{-1}) + f(\cdot) \\ A_{1i}x_{-1}(\tau_i-,c_{-1}) + \gamma_i \end{bmatrix}\right)(t)dt\right. \\ &\left.+ \sum_{i=-\infty}^{\infty} H_d(\tau_i)A_{1i}\left(G\begin{bmatrix} A_1(\cdot)x_{-1}(\cdot,c_{-1}) + f(\cdot) \\ A_{1i}x_{-1}(\cdot,c_{-1}) + \gamma_i \end{bmatrix}\right)(\tau_i-)\right]. \end{aligned} \tag{3.18}$$

Thus, under the condition (3.4), system (3.16) possesses an r-parameter set of solutions bounded on $\mathbb{R}$ in the form

$$x_1(t,c_1) = X_r(t)c_1 + \left(G\begin{bmatrix} A_1(\cdot)x_0(\cdot,c_0) \\ A_{1i}x(\tau_i-,c_0) \end{bmatrix}\right)(t), \tag{3.19}$$

where $(G[\substack{*\\*}])(t)$ is the generalized Green operator (2.7) of the problem of finding bounded solutions of system (3.16), and c_1 is an r-dimensional constant vector determined in the next stage of the process from the condition of solvability of the problem for $x_2(t)$.

If we continue this process, we prove (by induction) that the problem of determination of the coefficient $x_k(t)$ in the series (3.5) is reduced to the problem of finding solutions of the system

$$\begin{aligned} \dot{x}_k &= A(t)x_k + A_1(t)x_{k-1}, \quad t \notin \mathcal{T}, \\ \Delta x_k|_{t=\tau_i} &= A_{1i}x_{k-1}(\tau_i-), \quad i \in \mathbb{Z}, \; k = 1,2,\ldots, \end{aligned} \tag{3.20}$$

to Contents, page 1

bounded on the entire real axis. If the condition (3.4) is satisfied, then a solution of this problem bounded on $\mathbb{R}$ has the form

$$x_k(t) = x_k(t, c_k) = X_r(t)c_k + \left(G\left[\begin{matrix} A_1(\cdot)x_{k-1}(\cdot, c_{k-1}) \\ A_{1k}x_{k-1}(\tau_i -, c_{k-1}) \end{matrix} \right] \right)(t), \tag{3.21}$$

where $(G[{}^*_*])(t)$ is the generalized Green operator of the problem of finding bounded solutions of impulsive system (3.20) and the constant vector $c_k \in R^r$ is given by the formula

$$\begin{aligned} c_k = -B_0^+ \Bigg[\int_{-\infty}^{\infty} H_d(t)A_1(t)\left(G\left[\begin{matrix} A_1(\cdot)x_{k-1}(\cdot, c_{k-1}) \\ A_{1i}x_{k-1}(\tau_i -, c_{k-1}) \end{matrix} \right] \right)(t)dt \\ + \sum_{i=-\infty}^{\infty} H_d(\tau_i)A_{1i}\left(G\left[\begin{matrix} A_1(\cdot)x_{k-1}(\cdot, c_{k-1}) \\ A_{1i}x_{k-1}(\cdot, c_{k-1}) \end{matrix} \right] \right)(\tau_i -) \Bigg] \end{aligned} \tag{3.22}$$

(within an arbitrary vector constant $P_{B_0}c$, $c \in R^r$).

The fact that the series (3.5) is convergent can be proved by using the procedure of majorization. □

In the case where the number $r = \operatorname{rank} PP_D = \operatorname{rank}(I - Q)P_D$ of linear independent solutions of system (2.2) bounded on $\mathbb{R}$ is equal to the number $d = \operatorname{rank}[P_{D^*}Q] = \operatorname{rank}[P_{D^*}(I - P)]$, Theorem 3.1 yields the following assertion.

Corollary 3.2. *Suppose that the system* (3.1) *satisfies the conditions imposed above, and the homogeneous system* (2.2) *is e-dichotomous on* $\mathbb{R}_+$ *and* $\mathbb{R}_-$ *with projectors* P *and* Q*, respectively. Let nonhomogeneities* $f \in BC_\tau(\mathbb{R})$ *and* $\gamma_i \in \mathbb{R}^n$ *be given such that the condition* (2.4) *is not satisfied, and the generating system* (2.1) *does not have solutions bounded on the entire real axis. If condition*

$$\det B_0 \neq 0 \quad (r = d), \tag{3.23}$$

is satisfied, then the system (3.1) *possesses a unique solution bounded on* $\mathbb{R}$ *in the form of series* (3.5) *uniformly convergent for sufficiently small fixed* $\varepsilon \in (0, \varepsilon_*]$.

Proof. If $r = d$, then B_0 is a square matrix. Therefore, it follows from condition (3.4) that $P_{B_0} = P_{B_0^*} = 0$, which is equivalent to the condition (3.23). In this case, the constant vectors $c_k \in \mathbb{R}^r$ are uniquely determined from (3.22). The coefficients of the series (3.5) are also uniquely determined by (3.21), and, for all $f \in BC_\tau(\mathbb{R})$ and $\gamma_i \in \mathbb{R}^n$, the system (3.1) possesses a unique solution bounded on $\mathbb{R}$, which means that system (3.2) is e-dichotomous. □

We now illustrate the assertions proved above.

Example 3.3. Consider the impulsive system

$$\dot{x} = A(t)x + \varepsilon A_1(t)x + f(t), \quad t \in \mathbb{R} \setminus \mathcal{T},$$
$$\Delta x|_{t=\tau_i} = \gamma_i + \varepsilon A_{1i}x(\tau_i -), \quad \gamma_i = \begin{Bmatrix} \gamma_i^{(1)} \\ \gamma_i^{(2)} \\ \gamma_i^{(3)} \end{Bmatrix} \in \mathbb{R}^3, \ i \in \mathbb{Z}, \tag{3.24}$$

where

$$\begin{gathered} A(t) = \operatorname{diag}\{-\tanh t, \ -\tanh t, \tanh t\}, \\ f(t) = \operatorname{col}\{f_1(t), f_2(t), f_3(t)\} \in BC_{\mathcal{T}}(\mathbb{R}), \\ A_1(t) = \{a_{ij}(t)\}_{i,j=1}^{3} \in BC_{\mathcal{T}}(\mathbb{R}), \qquad A_{1i} = \{\tilde{a}_{ij}\}_{i,j=1}^{3}. \end{gathered} \tag{3.25}$$

The generating homogenous system (for $\varepsilon = 0$) has the form

$$\dot{x} = A(t)x, \qquad \Delta x|_{t=\tau_i} = 0 \tag{3.26}$$

and is e-dichotomous (as shown in [6]) on the semiaxes $\mathbb{R}_+$ and $\mathbb{R}_-$ with projectors $P = \operatorname{diag}\{1,1,0\}$ and $Q = \operatorname{diag}\{0,0,1\}$. The normal fundamental matrix of this system is

$$X(t) = \operatorname{diag}\left\{\frac{2}{e^t + e^{-t}}, \frac{2}{e^t + e^{-t}}, \frac{e^t + e^{-t}}{2}\right\}. \tag{3.27}$$

Thus, we have

$$\begin{gathered} D = 0, \qquad D^+ = 0, \qquad P_D = P_{D^*} = I_3, \\ r = \operatorname{rank} PP_D = 2, \qquad d = \operatorname{rank} P_{D^*}Q = 1, \\ X_r(t) = \begin{pmatrix} \dfrac{2}{e^t + e^{-t}} & 0 \\ 0 & \dfrac{2}{e^t + e^{-t}} \\ 0 & 0 \end{pmatrix}, \end{gathered} \tag{3.28}$$

$$H_d(t) = \left(0, 0, \frac{2}{e^t + e^{-t}}\right). \tag{3.29}$$

In order that the generating impulsive system (2.1) with the matrix $A(t)$ specified above has solutions bounded on the entire real axis, the nonhomogeneities $f(t) = \operatorname{col}\{f_1(t), f_2(t), f_3(t)\} \in BC_{\mathcal{T}}(\mathbb{R})$ and $\gamma_i = \operatorname{col}\{\gamma_i^{(1)}, \gamma_i^{(2)}, \gamma_i^{(3)}\} \in \mathbb{R}^3$ must satisfy condition (2.4). In this analyzed impulsive problem, this condition takes the form

$$\int_{-\infty}^{\infty} \frac{2 f_3(t)}{e^t + e^{-t}} dt + \sum_{i=-\infty}^{\infty} \frac{2}{e^{\tau_i} + e^{-\tau_i}} \gamma_i^{(3)} = 0, \quad \forall f_1(t), f_2(t) \in BC_{\mathcal{T}}(\mathbb{R}), \ \forall \gamma_i^{(1)}, \gamma_i^{(2)} \in \mathbb{R}. \tag{3.30}$$

to Contents, page 1

Let f_3 and $\gamma_i^{(3)}$ be given such that the condition (3.30) is not satisfied and the corresponding generating system (2.1) does not have solutions bounded on the entire real axis. The system (3.24) will be an e-trichotomous on $\mathbb{R}$ if the coefficients $a_{31}(t), a_{32}(t) \in BC_{\mathcal{T}}(\mathbb{R})$ of the perturbing matrix $A_1(t)$ and the coefficients $\tilde{a}_{31}, \tilde{a}_{32} \in \mathbb{R}$ of the perturbing matrix A_{1i} satisfy condition (3.4), that is, $P_{B_0^*} = 0$, where the matrix B_0 has the form

$$B_0 = \int_{-\infty}^{\infty} \left[\frac{a_{31}(t)}{(e^t + e^{-t})^2}, \frac{a_{32}(t)}{(e^t + e^{-t})^2} \right] dt + \sum_{i=-\infty}^{\infty} \left[\frac{\tilde{a}_{31}}{(e^{\tau_i -}+ e^{-\tau_i -})^2}, \frac{\tilde{a}_{32}}{(e^{\tau_i -} + e^{-\tau_i -})^2} \right]. \tag{3.31}$$

Therefore, if $a_{31}(t), a_{32}(t) \in BC_{\mathcal{T}}(\mathbb{R})$ and $\tilde{a}_{31}, \tilde{a}_{32} \in \mathbb{R}$ are such that at least one of the following inequalities

$$\begin{gathered} \int_{-\infty}^{\infty} \frac{a_{31}(t)}{(e^t + e^{-t})^2} dt + \sum_{i=-\infty}^{\infty} \frac{\tilde{a}_{31}}{(e^{\tau_i -} + e^{-\tau_i -})^2} \neq 0, \\ \int_{-\infty}^{\infty} \frac{a_{32}(t)}{(e^t + e^{-t})^2} dt + \sum_{i=-\infty}^{\infty} \frac{\tilde{a}_{32}}{(e^{\tau_i -} + e^{-\tau_i -})^2} \neq 0 \end{gathered} \tag{3.32}$$

is satisfied, then either the condition (3.4) or the equivalent condition $\operatorname{rank} B_0 = d = 1$ from Theorem 3.1 is satisfied and the system (3.2) is e-trichotomous on $\mathbb{R}$. In this case, the coefficients $a_{11}(t), a_{12}(t), a_{13}(t), a_{21}(t), a_{22}(t), a_{23}(t), a_{33}(t)$ are arbitrary functions from the space $BC_{\mathcal{T}}(\mathbb{R})$, and $\tilde{a}_{11}, \tilde{a}_{12}, \tilde{a}_{13}, \tilde{a}_{21}, \tilde{a}_{22}, \tilde{a}_{23}, \tilde{a}_{33}$ are arbitrary constants from $\mathbb{R}$. Moreover, for any

$$f(t) = \operatorname{col}\{ f_1(t), f_2(t), f_3(t) \} \in BC_{\mathcal{T}}(\mathbb{R}) \tag{3.33}$$

a solution of the system (3.24) bounded on $\mathbb{R}$ is given by the series (3.5) (within a constant from the null space $N(B_0)$, $\dim N(B_0) = r - \operatorname{rank} B_0 = 1$).

Another Perturbed Problem

In this part, we show that the problem of finding bounded solutions of nonhomogeneous system (2.1), in the case if the condition (2.4) is not satisfied, can be made solvable by introducing linear perturbations only to the pulsed conditions.

Therefore, we consider the weakly perturbed nonhomogeneous linear impulsive system in the form

$$\begin{gathered} \dot{x} = A(t)x + f(t), \quad t \in \mathbb{R} \setminus \mathcal{T},\ A, f \in BC_{\mathcal{T}}(\mathbb{R}), \\ \Delta x|_{t=\tau_i} = \gamma_i + \varepsilon A_{1i} x(\tau_i -), \quad \gamma_i \in \mathbb{R}^n,\ i \in \mathbb{Z}, \end{gathered} \tag{3.34}$$

where A_{1i} are $n \times n$ constant matrices. For $\varepsilon = 0$, we obtain the generating system (2.1). We assume that this generating system does not have solutions bounded on the entire real axis, which means that the condition of solvability (2.4) is not satisfied (for some nonhomogeneities $f \in BC_{\mathcal{T}}(\mathbb{R})$ and $\gamma_i \in \mathbb{R}^n$). Let us show that it is possible to make this problem solvable by adding linear perturbation only to the pulsed conditions. In the case, if

to Contents, page 1

this is possible, it is necessary to determine perturbations A_{1i} for which the corresponding homogeneous system

$$\begin{aligned} \dot{x} &= A(t)x, \quad t \in \mathbb{R} \setminus \mathcal{T}, \\ \Delta x|_{t=\tau_i} &= \varepsilon A_{1i} x(\tau_i -), \quad i \in \mathbb{Z}, \end{aligned} \tag{3.35}$$

turns into the system e-trichotomous or e-dichotomous on the entire real axis.

This problem can be solved with help of the $d \times r$ matrix

$$B_0 = \sum_{i=-\infty}^{\infty} H_d(\tau_i) A_{1i} X_r(\tau_i -) \tag{3.36}$$

constructed with the coefficients from the impulsive system (3.34).

By using Theorem 3.1, we seek a solution in the form of the series (3.5). Thus, we have the following corollary.

Corollary 3.4. *Suppose that the system* (3.34) *satisfies the conditions imposed above and the generating homogeneous system* (2.2) *is e-dichotomous on* $\mathbb{R}_+$ *and* $\mathbb{R}_-$ *with projectors* P *and* Q*, respectively. Let nonhomogeneities* $f \in BC_{\mathcal{T}}(\mathbb{R})$ *and* $\gamma_i \in \mathbb{R}^n$ *be given such that the condition* (2.4) *is not satisfied, and the generating system* (2.1) *does not have solutions bounded on the entire real axis. If the condition* (3.4) *is satisfied, then the system* (3.35) *is e-trichotomous on* $\mathbb{R}$*, and the system* (3.34) *possesses at least one solution bounded on* $\mathbb{R}$ *in the form of series* (3.5) *uniformly convergent for sufficiently small fixed* $\varepsilon \in (0, \varepsilon_*]$.

Acknowledgments

This research was supported by Grants 1/0771/08 of the Grant Agency of Slovak Republic (VEGA) and project APVV-0700-07 of Slovak Research and Development Agency.

References

[1] A. M. Samoilenko and N. A. Perestyuk, *Impulsive Differential Equations*, Vyshcha Shkola, Kiev, Ukraine, 1974.

[2] S. Schwabik, M. Tvrdý, and O. Vejvoda, *Differential and Integral Equations*, Academia, Prague, Czech Republic, 1979.

[3] A. A. Boichuk, N. A. Perestyuk, and A. M. Samoilenko, "Periodic solutions of impulse differential systems in critical cases," *Differents. Uravn.*, vol. 27, no. 9, pp. 1516–1521, 1991.

[4] R. J. Sacker, "The splitting index for linear differential systems," *Journal of Differential Equations*, vol. 33, no. 3, pp. 368–405, 1979.

[5] K. J. Palmer, "Exponential dichotomies and transversal homoclinic points," *Journal of Differential Equations*, vol. 55, no. 2, pp. 225–256, 1984.

[6] A. M. Samoilenko, A. A. Boichuk, and An. A. Boichuk, "Solutions, bounded on the whole axis, of linear weakly perturbed systems," *Ukrainian Mathematical Journal*, vol. 54, no. 11, pp. 1517–1530, 2002.

[7] A. A. Boichuk, "Solutions of weakly nonlinear differential equations bounded on the whole line," *Nonlinear Oscillations*, vol. 2, no. 1, pp. 3–10, 1999.

[8] A. Boichuk and A. Pokutnyi, "Bounded solutions of linear perturbed differential equations in a Banach space," *Tatra Mountains Mathematical Publications*, vol. 39, pp. 1–12, 2007.

[9] F. Battelli and M. Fečkan, "Chaos in singular impulsive O.D.E.," *Nonlinear Analysis*, vol. 28, no. 4, pp. 655–671, 1997.

to Contents, page 1

[10] A. A. Boichuk and A. M. Samoilenko, *Generalized Inverse Operators and Fredholm Boundary-Value Problems*, Koninklijke Brill NV, Utrecht, Mass, USA, 2004.

[11] A. A. Boichuk, J. Diblík, D. Khusainov, and M. Růžičková, "Fredholm's boundary-value problems for differential systems with a single delay," *Nonlinear Analysis*, vol. 72, no. 5, pp. 2251–2258, 2010.

[12] A. A. Boichuk, J. Diblík, D. Khusainov, and M. Růžičková, "Boundary-value problems for weakly nonlinear delay differential systems," *Abstract and Applied Analysis*, vol. 2011, Article ID 631412, 19 pages, 2011.

[13] A. Boichuk, M. Langerová, and J. Škoríková, "Solutions of linear impulsive differential systems bounded on the entire real axis," *Advances in Difference Equations*, vol. 2010, Article ID 494379, 10 pages, 2010.

[14] M. I. Vishik and L. A. Lyusternik, "The solution of some perturbation problems for matrices and selfadjoint differential equations," *Uspekhi Matematicheskikh Nauk*, vol. 15, no. 3, pp. 3–80, 1960.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 346745, 9 pages
doi:10.1155/2011/346745

Research Article

Existence of Nonoscillatory Solutions of First-Order Neutral Differential Equations

Božena Dorociaková, Anna Najmanová, and Rudolf Olach

Department of Mathematics, University of Žilina, 010 26 Žilina, Slovakia

Correspondence should be addressed to Božena Dorociaková, bozena.dorociakova@fstroj.uniza.sk

Received 4 January 2011; Revised 4 March 2011; Accepted 27 April 2011

Academic Editor: Josef Diblík

This paper contains some sufficient conditions for the existence of positive solutions which are bounded below and above by positive functions for the first-order nonlinear neutral differential equations. These equations can also support the existence of positive solutions approaching zero at infinity

1. Introduction

This paper is concerned with the existence of a positive solution of the neutral differential equations of the form

$$\frac{d}{dt}[x(t) - a(t)x(t-\tau)] = p(t)f(x(t-\sigma)), \quad t \geq t_0, \tag{1.1}$$

where $\tau > 0$, $\sigma \geq 0$, $a \in C([t_0,\infty),(0,\infty))$, $p \in C(R,(0,\infty))$, $f \in C(R,R)$, f is nondecreasing function, and $xf(x) > 0$, $x \neq 0$.

By a solution of (1.1) we mean a function $x \in C([t_1 - m,\infty), R)$, $m = \max\{\tau,\sigma\}$, for some $t_1 \geq t_0$, such that $x(t) - a(t)x(t-\tau)$ is continuously differentiable on $[t_1,\infty)$ and such that (1.1) is satisfied for $t \geq t_1$.

The problem of the existence of solutions of neutral differential equations has been studied by several authors in the recent years. For related results we refer the reader to [1–11] and the references cited therein. However there is no conception which guarantees the existence of positive solutions which are bounded below and above by positive functions. In this paper we have presented some conception. The method also supports the existence of positive solutions approaching zero at infinity.

to Contents, page 1

As much as we know, for (1.1) in the literature, there is no result for the existence of solutions which are bounded by positive functions. Only the existence of solutions which are bounded by constants is treated, for example, in [6, 10, 11]. It seems that conditions of theorems are rather complicated, but cannot be simpler due to Corollaries 2.3, 2.6, and 3.2.

The following fixed point theorem will be used to prove the main results in the next section.

Lemma 1.1 ([see [6, 10] Krasnoselskii's fixed point theorem]). *Let X be a Banach space, let Ω be a bounded closed convex subset of X, and let S_1, S_2 be maps of Ω into X such that $S_1x + S_2y \in \Omega$ for every pair $x, y \in \Omega$. If S_1 is contractive and S_2 is completely continuous, then the equation*

$$S_1x + S_2x = x \tag{1.2}$$

has a solution in Ω.

2. The Existence of Positive Solution

In this section we will consider the existence of a positive solution for (1.1). The next theorem gives us the sufficient conditions for the existence of a positive solution which is bounded by two positive functions.

Theorem 2.1. *Suppose that there exist bounded functions $u, v \in C^1([t_0, \infty), (0, \infty))$, constant $c > 0$ and $t_1 \geq t_0 + m$ such that*

$$u(t) \leq v(t), \quad t \geq t_0, \tag{2.1}$$

$$v(t) - v(t_1) - u(t) + u(t_1) \geq 0, \quad t_0 \leq t \leq t_1, \tag{2.2}$$

$$\begin{aligned} &\frac{1}{u(t-\tau)}\left(u(t) + \int_t^\infty p(s) f(v(s-\sigma))ds\right) \leq a(t) \\ &\leq \frac{1}{v(t-\tau)}\left(v(t) + \int_t^\infty p(s) f(u(s-\sigma))ds\right) \leq c < 1, \quad t \geq t_1. \end{aligned} \tag{2.3}$$

Then (1.1) *has a positive solution which is bounded by functions u, v.*

Proof. Let $C([t_0, \infty), R)$ be the set of all continuous bounded functions with the norm $\|x\| = \sup_{t\geq t_0}|x(t)|$. Then $C([t_0, \infty), R)$ is a Banach space. We define a closed, bounded, and convex subset Ω of $C([t_0, \infty), R)$ as follows:

$$\Omega = \{x = x(t) \in C([t_0, \infty), R) : u(t) \leq x(t) \leq v(t),\ t \geq t_0\}. \tag{2.4}$$

to Contents, page 1

We now define two maps S_1 and $S_2 : \Omega \longrightarrow C([t_0, \infty), R)$ as follows:

$$(S_1x)(t) = \begin{cases} a(t)x(t-\tau), & t \geq t_1, \\ (S_1x)(t_1), & t_0 \leq t \leq t_1, \end{cases}$$
$$(S_2x)(t) = \begin{cases} -\displaystyle\int_t^\infty p(s)f(x(s-\sigma))ds, & t \geq t_1, \\ (S_2x)(t_1) + v(t) - v(t_1), & t_0 \leq t \leq t_1. \end{cases} \tag{2.5}$$

We will show that for any $x, y \in \Omega$ we have $S_1x + S_2y \in \Omega$. For every $x, y \in \Omega$ and $t \geq t_1$, we obtain

$$(S_1x)(t) + \left(S_2y\right)(t) \leq a(t)v(t-\tau) - \int_t^\infty p(s)f(u(s-\sigma))ds \leq v(t). \tag{2.6}$$

For $t \in [t_0, t_1]$, we have

$$\begin{aligned}(S_1x)(t) + \left(S_2y\right)(t) &= (S_1x)(t_1) + \left(S_2y\right)(t_1) + v(t) - v(t_1) \\ &\leq v(t_1) + v(t) - v(t_1) = v(t).\end{aligned} \tag{2.7}$$

Furthermore, for $t \geq t_1$, we get

$$(S_1x)(t) + \left(S_2y\right)(t) \geq a(t)u(t-\tau) - \int_t^\infty p(s)f(v(s-\sigma))ds \geq u(t). \tag{2.8}$$

Let $t \in [t_0, t_1]$. With regard to (2.2), we get

$$v(t) - v(t_1) + u(t_1) \geq u(t), \quad t_0 \leq t \leq t_1. \tag{2.9}$$

Then for $t \in [t_0, t_1]$ and any $x, y \in \Omega$, we obtain

$$\begin{aligned}(S_1x)(t) + \left(S_2y\right)(t) &= (S_1x)(t_1) + \left(S_2y\right)(t_1) + v(t) - v(t_1) \\ &\geq u(t_1) + v(t) - v(t_1) \geq u(t).\end{aligned} \tag{2.10}$$

Thus we have proved that $S_1x + S_2y \in \Omega$ for any $x, y \in \Omega$.

We will show that S_1 is a contraction mapping on Ω. For $x, y \in \Omega$ and $t \geq t_1$ we have

$$\left|(S_1x)(t) - (S_1y)(t)\right| = \left|a(t)\right|\left|x(t-\tau) - y(t-\tau)\right| \leq c\left\|x - y\right\|. \tag{2.11}$$

This implies that

$$\left\|S_1x - S_1y\right\| \leq c\left\|x - y\right\|. \tag{2.12}$$

Also for $t \in [t_0, t_1]$, the previous inequality is valid. We conclude that S_1 is a contraction mapping on Ω.

We now show that S_2 is completely continuous. First we will show that S_2 is continuous. Let $x_k = x_k(t) \in \Omega$ be such that $x_k(t) \to x(t)$ as $k \to \infty$. Because Ω is closed, $x = x(t) \in \Omega$. For $t \geq t_1$ we have

$$\begin{aligned} |(S_2x_k)(t) - (S_2x)(t)| &\leq \left| \int_t^\infty p(s) \left[f(x_k(s-\sigma)) - f(x(s-\sigma)) \right] ds \right| \\ &\leq \int_{t_1}^\infty p(s) \left| f(x_k(s-\sigma)) - f(x(s-\sigma)) \right| ds. \end{aligned} \tag{2.13}$$

According to (2.8), we get

$$\int_{t_1}^\infty p(s) f(v(s-\sigma)) ds < \infty. \tag{2.14}$$

Since $|f(x_k(s-\sigma)) - f(x(s-\sigma))| \to 0$ as $k \to \infty$, by applying the Lebesgue dominated convergence theorem, we obtain

$$\lim_{k\to\infty} \|(S_2x_k)(t) - (S_2x)(t)\| = 0. \tag{2.15}$$

This means that S_2 is continuous.

We now show that $S_2\Omega$ is relatively compact. It is sufficient to show by the Arzela-Ascoli theorem that the family of functions $\{S_2x : x \in \Omega\}$ is uniformly bounded and equicontinuous on $[t_0, \infty)$. The uniform boundedness follows from the definition of Ω. For the equicontinuity we only need to show, according to Levitans result [7], that for any given $\varepsilon > 0$ the interval $[t_0, \infty)$ can be decomposed into finite subintervals in such a way that on each subinterval all functions of the family have a change of amplitude less than ε. Then with regard to condition (2.14), for $x \in \Omega$ and any $\varepsilon > 0$, we take $t^* \geq t_1$ large enough so that

$$\int_{t^*}^\infty p(s) f(x(s-\sigma)) ds < \frac{\varepsilon}{2}. \tag{2.16}$$

Then, for $x \in \Omega$, $T_2 > T_1 \geq t^*$, we have

$$\begin{aligned} |(S_2x)(T_2) - (S_2x)(T_1)| &\leq \int_{T_2}^\infty p(s) f(x(s-\sigma)) ds \\ &\quad + \int_{T_1}^\infty p(s) f(x(s-\sigma)) ds < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned} \tag{2.17}$$

to Contents, page 1

For $x \in \Omega$ and $t_1 \leq T_1 < T_2 \leq t^*$, we get

$$\begin{aligned} |(S_2x)(T_2) - (S_2x)(T_1)| &\leq \int_{T_1}^{T_2} p(s)f(x(s-\sigma))ds \\ &\leq \max_{t_1 \leq s \leq t^*}\{p(s)f(x(s-\sigma))\}(T_2 - T_1). \end{aligned} \tag{2.18}$$

Thus there exists $\delta_1 = \varepsilon/M$, where $M = \max_{t_1 \leq s \leq t^*}\{p(s)f(x(s-\sigma))\}$, such that

$$|(S_2x)(T_2) - (S_2x)(T_1)| < \varepsilon \quad \text{if } 0 < T_2 - T_1 < \delta_1. \tag{2.19}$$

Finally for any $x \in \Omega$, $t_0 \leq T_1 < T_2 \leq t_1$, there exists a $\delta_2 > 0$ such that

$$\begin{aligned} |(S_2x)(T_2) - (S_2x)(T_1)| = |v(T_1) - v(T_2)| &= \left|\int_{T_1}^{T_2} v'(s)ds\right| \\ &\leq \max_{t_0 \leq s \leq t_1}\{|v'(s)|\}(T_2 - T_1) < \varepsilon \quad \text{if } 0 < T_2 - T_1 < \delta_2. \end{aligned} \tag{2.20}$$

Then $\{S_2x : x \in \Omega\}$ is uniformly bounded and equicontinuous on $[t_0, \infty)$, and hence $S_2\Omega$ is relatively compact subset of $C([t_0, \infty), R)$. By Lemma 1.1 there is an $x_0 \in \Omega$ such that $S_1x_0 + S_2x_0 = x_0$. We conclude that $x_0(t)$ is a positive solution of (1.1). The proof is complete. □

Corollary 2.2. *Suppose that there exist functions $u, v \in C^1([t_0, \infty), (0, \infty))$, constant $c > 0$ and $t_1 \geq t_0 + m$ such that* (2.1), (2.3) *hold and*

$$v'(t) - u'(t) \leq 0, \quad t_0 \leq t \leq t_1. \tag{2.21}$$

Then (1.1) *has a positive solution which is bounded by the functions u, v.*

Proof. We only need to prove that condition (2.21) implies (2.2). Let $t \in [t_0, t_1]$ and set

$$H(t) = v(t) - v(t_1) - u(t) + u(t_1). \tag{2.22}$$

Then with regard to (2.21), it follows that

$$H'(t) = v'(t) - u'(t) \leq 0, \quad t_0 \leq t \leq t_1. \tag{2.23}$$

Since $H(t_1) = 0$ and $H'(t) \leq 0$ for $t \in [t_0, t_1]$, this implies that

$$H(t) = v(t) - v(t_1) - u(t) + u(t_1) \geq 0, \quad t_0 \leq t \leq t_1. \tag{2.24}$$

Thus all conditions of Theorem 2.1 are satisfied. □

Corollary 2.3. *Suppose that there exists a function* $v \in C^1([t_0, \infty), (0, \infty))$, *constant* $c > 0$ *and* $t_1 \geq t_0 + m$ *such that*

$$a(t) = \frac{1}{v(t-\tau)}\left(v(t) + \int_t^\infty p(s) f(v(s-\sigma))ds\right) \leq c < 1, \quad t \geq t_1. \tag{2.25}$$

Then (1.1) *has a solution* $x(t) = v(t)$, $t \geq t_1$.

Proof. We put $u(t) = v(t)$ and apply Theorem 2.1. □

Theorem 2.4. *Suppose that there exist functions* $u, v \in C^1([t_0, \infty), (0, \infty))$, *constant* $c > 0$ *and* $t_1 \geq t_0 + m$ *such that* (2.1), (2.2), *and* (2.3) *hold and*

$$\lim_{t \to \infty} v(t) = 0. \tag{2.26}$$

Then (1.1) *has a positive solution which is bounded by the functions* u, v *and tends to zero.*

Proof. The proof is similar to that of Theorem 2.1 and we omit it. □

Corollary 2.5. *Suppose that there exist functions* $u, v \in C^1([t_0, \infty), (0, \infty))$, *constant* $c > 0$ *and* $t_1 \geq t_0 + m$ *such that* (2.1), (2.3), (2.21), *and* (2.26) *hold. Then* (1.1) *has a positive solution which is bounded by the functions* u, v *and tends to zero.*

Proof. The proof is similar to that of Corollary 2.2, and we omitted it. □

Corollary 2.6. *Suppose that there exists a function* $v \in C^1([t_0, \infty), (0, \infty))$, *constant* $c > 0$ *and* $t_1 \geq t_0 + m$ *such that* (2.25), (2.26) *hold. Then* (1.1) *has a solution* $x(t) = v(t)$, $t \geq t_1$ *which tends to zero.*

Proof. We put $u(t) = v(t)$ and apply Theorem 2.4. □

3. Applications and Examples

In this section we give some applications of the theorems above.

Theorem 3.1. *Suppose that*

$$\int_{t_0}^\infty p(t)dt = \infty, \tag{3.1}$$

to Contents, page 1

$0 < k_1 \le k_2$ *and there exist constants* $c > 0$, $\gamma \ge 0$, $t_1 \ge t_0 + m$ *such that*

$$\frac{k_1}{k_2}\exp\left((k_2-k_1)\int_{t_0-\gamma}^{t_0} p(t)dt\right) \ge 1, \tag{3.2}$$

$$\begin{aligned}
&\exp\left(-k_2\int_{t-\tau}^{t} p(s)ds\right) + \exp\left(k_2\int_{t_0-\gamma}^{t-\tau} p(s)ds\right)\\
&\quad\times\int_t^{\infty} p(s)f\left(\exp\left(-k_1\int_{t_0-\gamma}^{s-\sigma} p(\xi)d\xi\right)\right)ds \le a(t)\\
&\le \exp\left(-k_1\int_{t-\tau}^{t} p(s)ds\right) + \exp\left(k_1\int_{t_0-\gamma}^{t-\tau} p(s)ds\right)\\
&\quad\times\int_t^{\infty} p(s)f\left(\exp\left(-k_2\int_{t_0-\gamma}^{s-\sigma} p(\xi)d\xi\right)\right)ds \le c < 1, \quad t \ge t_1.
\end{aligned} \tag{3.3}$$

Then (1.1) *has a positive solution which tends to zero.*

Proof. We set

$$u(t) = \exp\left(-k_2\int_{t_0-\gamma}^{t} p(s)ds\right), \quad v(t) = \exp\left(-k_1\int_{t_0-\gamma}^{t} p(s)ds\right), \quad t \ge t_0. \tag{3.4}$$

We will show that the conditions of Corollary 2.5 are satisfied. With regard to (2.21), for $t \in [t_0, t_1]$, we get

$$\begin{aligned}
v'(t) - u'(t) &= -k_1 p(t)v(t) + k_2 p(t)u(t)\\
&= p(t)v(t)\left[-k_1 + k_2 u(t)\exp\left(k_1\int_{t_0-\gamma}^{t} p(s)ds\right)\right]\\
&= p(t)v(t)\left[-k_1 + k_2\exp\left((k_1-k_2)\int_{t_0-\gamma}^{t} p(s)ds\right)\right]\\
&\le p(t)v(t)\left[-k_1 + k_2\exp\left((k_1-k_2)\int_{t_0-\gamma}^{t_0} p(s)ds\right)\right] \le 0.
\end{aligned} \tag{3.5}$$

Other conditions of Corollary 2.5 are also satisfied. The proof is complete. □

to Contents, page 1

Corollary 3.2. *Suppose that* $k > 0,\ c > 0,\ t_1 \geq t_0 + m$, (3.1) *holds, and*

$$\begin{aligned} a(t) = \exp\left(-k\int_{t-\tau}^{t} p(s)ds\right) + \exp\left(k\int_{t_0}^{t-\tau} p(s)ds\right) \\ \times \int_{t}^{\infty} p(s) f\left(\exp\left(-k\int_{t_0}^{s-\sigma} p(\xi)d\xi\right)\right)ds \leq c < 1, \quad t \geq t_1. \end{aligned} \tag{3.6}$$

Then (1.1) *has a solution*

$$x(t) = \exp\left(-k\int_{t_0}^{t} p(s)ds\right), \quad t \geq t_1, \tag{3.7}$$

which tends to zero.

Proof. We put $k_1 = k_2 = k,\ \gamma = 0$ and apply Theorem 3.1. □

Example 3.3. Consider the nonlinear neutral differential equation

$$[x(t) - a(t)x(t-2)]' = px^3(t-1), \quad t \geq t_0, \tag{3.8}$$

where $p \in (0, \infty)$. We will show that the conditions of Theorem 3.1 are satisfied. Condition (3.1) obviously holds and (3.2) has a form

$$\frac{k_1}{k_2}\exp((k_2 - k_1)p\gamma) \geq 1, \tag{3.9}$$

$0 < k_1 \leq k_2,\ \gamma \geq 0$. For function $a(t)$, we obtain

$$\begin{aligned} &\exp(-2pk_2) + \frac{1}{3k_1}\exp(p[k_2(\gamma - t_0 - 2) - 3k_1(\gamma - t_0 - 1) + (k_2 - 3k_1)t]) \\ &\quad \leq a(t) \leq \exp(-2pk_1) \\ &\quad + \frac{1}{3k_2}\exp(p[k_1(\gamma - t_0 - 2) - 3k_2(\gamma - t_0 - 1) + (k_1 - 3k_2)t]), \quad t \geq t_0. \end{aligned} \tag{3.10}$$

For $p = 1,\ k_1 = 1,\ k_2 = 2,\ \gamma = 1,\ t_0 = 1$, condition (3.9) is satisfied and

$$e^{-4} + \frac{1}{3e}e^{-t} \leq a(t) \leq e^{-2} + \frac{e^4}{6}e^{-5t}, \quad t \geq t_1 \geq 3. \tag{3.11}$$

If the function $a(t)$ satisfies (3.11), then (3.8) has a solution which is bounded by the functions $u(t) = \exp(-2t),\ v(t) = \exp(-t),\ t \geq 3$.

to Contents, page 1

For example if $p = 1, k_1 = k_2 = 1.5, \gamma = 1, t_0 = 1$, from (3.11) we obtain

$$a(t) = e^{-3} + \frac{e^{1.5}}{4.5} e^{-3t}, \tag{3.12}$$

and the equation

$$\left[x(t) - \left(e^{-3} + \frac{e^{1.5}}{4.5} e^{-3t}\right) x(t-2)\right]' = x^3(t-1), \quad t \geq 3, \tag{3.13}$$

has the solution $x(t) = \exp(-1.5t)$ which is bounded by the function $u(t)$ and $v(t)$.

Acknowledgments

The research was supported by the Grant 1/0090/09 and the Grant 1/1260/12 of the Scientific Grant Agency of the Ministry of Education of the Slovak Republic and Project APVV-0700-07 of the Slovak Research and Development Agency.

References

[1] A. Boichuk, J. Diblík, D. Khusainov, and M. Růžičková, "Fredholm's boundary-value problems for differential systems with a single delay," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 72, no. 5, pp. 2251–2258, 2010.

[2] J. Diblík, "Positive and oscillating solutions of differential equations with delay in critical case," *Journal of Computational and Applied Mathematics*, vol. 88, no. 1, pp. 185–202, 1998, Positive solutions of nonlinear problem.

[3] J. Diblík and M. Kúdelčíková, "Two classes of asymptotically different positive solutions of the equation $\dot{y}(t) = -f(t, y_t)$," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 70, no. 10, pp. 3702–3714, 2009.

[4] J. Diblík, Z. Svoboda, and Z. Šmarda, "Retract principle for neutral functional differential equations," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 71, no. 12, pp. e1393–e1400, 2009.

[5] J. Diblík and M. Růžičková, "Existence of positive solutions of a singular initial problem for a nonlinear system of differential equations," *The Rocky Mountain Journal of Mathematics*, vol. 3, pp. 923–944, 2004.

[6] L. H. Erbe, Q. Kong, and B. G. Zhang, *Oscillation Theory for Functional-Differential Equations*, vol. 190 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 1995.

[7] B. M. Levitan, "Some questions of the theory of almost periodic functions. I," *Uspekhi Matematicheskikh Nauk*, vol. 2, no. 5, pp. 133–192, 1947.

[8] X. Lin, "Oscillation of second-order nonlinear neutral differential equations," *Journal of Mathematical Analysis and Applications*, vol. 309, no. 2, pp. 442–452, 2005.

[9] X. Wang and L. Liao, "Asymptotic behavior of solutions of neutral differential equations with positive and negative coefficients," *Journal of Mathematical Analysis and Applications*, vol. 279, no. 1, pp. 326–338, 2003.

[10] Y. Zhou, "Existence for nonoscillatory solutions of second-order nonlinear differential equations," *Journal of Mathematical Analysis and Applications*, vol. 331, no. 1, pp. 91–96, 2007.

[11] Y. Zhou and B. G. Zhang, "Existence of nonoscillatory solutions of higher-order neutral differential equations with positive and negative coefficients," *Applied Mathematics Letters*, vol. 15, no. 7, pp. 867–874, 2002.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 408525, 20 pages
doi:10.1155/2011/408525

Research Article

Existence of Oscillatory Solutions of Singular Nonlinear Differential Equations

Irena Rachůnková, Lukáš Rachůnek, and Jan Tomeček

Department of Mathematics, Faculty of Science, Palacký University, 17. listopadu 12, 771 46 Olomouc, Czech Republic

Correspondence should be addressed to Irena Rachůnková, irena.rachunkova@upol.cz

Received 10 October 2010; Revised 25 February 2011; Accepted 23 March 2011

Academic Editor: Yuri V. Rogovchenko

Asymptotic properties of solutions of the singular differential equation $(p(t)u'(t))' = p(t)f(u(t))$ are described. Here, f is Lipschitz continuous on $\mathbb{R}$ and has at least two zeros 0 and $L > 0$. The function p is continuous on $[0, \infty)$ and has a positive continuous derivative on $(0, \infty)$ and $p(0) = 0$. Further conditions for f and p under which the equation has oscillatory solutions converging to 0 are given.

1. Introduction

For $k \in \mathbb{N}$, $k > 1$, and $L \in (0, \infty)$, consider the equation

$$u'' + \frac{k-1}{t}u' = f(u), \quad t \in (0, \infty), \tag{1.1}$$

where

$$f \in \mathrm{Lip}_{\mathrm{loc}}(\mathbb{R}), \qquad f(0) = f(L) = 0, \qquad f(x) < 0, \quad x \in (0, L), \tag{1.2}$$

$$\exists \overline{B} \in (-\infty, 0) : f(x) > 0, \quad x \in \left[\overline{B}, 0\right). \tag{1.3}$$

Let us put

$$F(x) = -\int_0^x f(z)\mathrm{d}z \quad \text{for } x \in \mathbb{R}. \tag{1.4}$$

to Contents, page 1

Moreover, we assume that f fulfils

$$F\left(\overline{B}\right) = F(L), \tag{1.5}$$

and denote

$$L_0 = \inf\left\{x < \overline{B} : f(x) > 0\right\} \geq -\infty. \tag{1.6}$$

Due to (1.2)–(1.4), we see that $F \in C^1(\mathbb{R})$ is decreasing and positive on $(L_0, 0)$ and increasing and positive on $(0, L]$.

Equation (1.1) arises in many areas. For example, in the study of phase transitions of Van der Waals fluids [1–3], in population genetics, where it serves as a model for the spatial distribution of the genetic composition of a population [4, 5], in the homogenous nucleation theory [6], and in relativistic cosmology for description of particles which can be treated as domains in the universe [7], in the nonlinear field theory, in particular, when describing bubbles generated by scalar fields of the Higgs type in the Minkowski spaces [8]. Numerical simulations of solutions of (1.1), where f is a polynomial with three zeros, have been presented in [9–11]. Close problems about the existence of positive solutions can be found in [12–14].

In this paper, we investigate a generalization of (1.1) of the form

$$\left(p(t)u'\right)' = p(t)f(u), \quad t \in (0, \infty), \tag{1.7}$$

where f satisfies (1.2)–(1.5) and p fulfils

$$p \in C[0, \infty) \cap C^1(0, \infty), \quad p(0) = 0, \tag{1.8}$$

$$p'(t) > 0,\ t \in (0, \infty), \quad \lim_{t \to \infty} \frac{p'(t)}{p(t)} = 0. \tag{1.9}$$

Equation (1.7) is singular in the sense that $p(0) = 0$. If $p(t) = t^{k-1}$, with $k > 1$, then p satisfies (1.8), (1.9), and (1.7) is equal to (1.1).

Definition 1.1. A function $u \in C^1[0, \infty) \cap C^2(0, \infty)$ which satisfies (1.7) for all $t \in (0, \infty)$ is called a *solution* of (1.7).

Consider a solution u of (1.7). Since $u \in C^1[0, \infty)$, we have $u(0), u'(0) \in \mathbb{R}$ and the assumption, $p(0) = 0$ yields $p(0)u'(0) = 0$. We can find $M > 0$ and $\delta > 0$ such that $|f(u(t))| \leq M$ for $t \in (0, \delta)$. Integrating (1.7), we get

$$\left|u'(t)\right| = \left|\frac{1}{p(t)} \int_0^t p(s)f(u(s))\mathrm{d}s\right| \leq \frac{M}{p(t)} \int_0^t p(s)\mathrm{d}s \leq Mt, \quad t \in (0, \delta). \tag{1.10}$$

to Contents, page 1

Consequently, the condition

$$u'(0) = 0 \tag{1.11}$$

is necessary for each solution of (1.7). Denote

$$u_{\sup} = \sup\{u(t) : t \in [0, \infty)\}. \tag{1.12}$$

Definition 1.2. Let u be a solution of (1.7). If $u_{\sup} < L$, then u is called a *damped* solution.

If a solution u of (1.7) satisfies $u_{\sup} = L$ or $u_{\sup} > L$, then we call u a bounding homoclinic solution or an escape solution. These three types of solutions have been investigated in [15–18]. Here, we continue the investigation of the existence and asymptotic properties of damped solutions. Due to (1.11) and Definition 1.2, it is reasonable to study solutions of (1.7) satisfying the initial conditions

$$u(0) = u_0 \in (L_0, L], \qquad u'(0) = 0. \tag{1.13}$$

Note that if $u_0 > L$, then a solution u of the problem (1.7), (1.13) satisfies $u_{\sup} > L$, and consequently u is not a damped solution. Assume that $L_0 > -\infty$, then $f(L_0) = 0$, and if we put $u_0 = L_0$, a solution u of (1.7), (1.13) is a constant function equal to L_0 on $[0, \infty)$. Since we impose no sign assumption on $f(x)$ for $x < L_0$, we do not consider the case $u_0 < L_0$. In fact, the choice of u_0 between two zeros L_0 and 0 of f has been motivated by some hydrodynamical model in [11].

A lot of papers are devoted to oscillatory solutions of nonlinear differential equations. Wong [19] published an account on a nonlinear oscillation problem originated from earlier works of Atkinson and Nehari. Wong's paper is concerned with the study of oscillatory behaviour of second-order Emden-Fowler equations

$$y''(x) + a(x)\left|y(x)\right|^{\gamma-1} y(x) = 0, \quad \gamma > 0, \tag{1.14}$$

where a is nonnegative and absolutely continuous on $(0, \infty)$. Both superlinear case ($\gamma > 1$) and sublinear case ($\gamma \in (0,1)$) are discussed, and conditions for the function a giving oscillatory or nonoscillatory solutions of (1.14) are presented; see also [20]. Further extensions of these results have been proved for more general differential equations. For example, Wong and Agarwal [21] or Li [22] worked with the equation

$$\left(a(t)\left(y'(t)\right)^{\sigma}\right)' + q(t) f\left(y(t)\right) = 0, \tag{1.15}$$

where $\sigma > 0$ is a positive quotient of odd integers, $a \in C^1(\mathbb{R})$ is positive, $q \in C(\mathbb{R})$, $f \in C^1(\mathbb{R})$, $xf(x) > 0$, $f'(x) \geq 0$ for all $x \neq 0$. Kulenović and Ljubović [23] investigated an equation

$$\left(r(t) g\left(y'(t)\right)\right)' + p(t) f\left(y(t)\right) = 0, \tag{1.16}$$

to Contents, page 1

where $g(u)/u \leq m$, $f(u)/u \geq k > 0$, or $f'(u) \geq k$ for all $u \neq 0$. The investigation of oscillatory and nonoscillatory solutions has been also realized in the class of quasilinear equations. We refer to the paper [24] by Ho, dealing with the equation

$$\left(t^{n-1}\Phi_p(u')\right)' + t^{n-1}\sum_{i=1}^{N}\alpha_i t^{\beta_i}\Phi_{q_i}(u) = 0, \tag{1.17}$$

where $1 < p < n$, $\alpha_i > 0$, $\beta_i \geq -p$, $q_i > p-1$, $i = 1, \ldots, N$, $\Phi_p(y) = |y|^{p-2}y$.

Oscillation results for the equation

$$\left(a(t)\Phi_p(x')\right)' + b(t)\Phi_q(x) = 0, \tag{1.18}$$

where $a, b \in C([0,\infty))$ are positive, can be found in [25]. We can see that the nonlinearity $f(y) = |y|^{\gamma-1}y$ in (1.14) is an increasing function on $\mathbb{R}$ having a unique zero at $y = 0$.

Nonlinearities in all the other (1.15)–(1.18) have similar globally monotonous behaviour. We want to emphasize that, in contrast to the above papers, the nonlinearity f in our (1.7) needs not be globally monotonous. Moreover, we deal with solutions of (1.7) starting at a singular point $t = 0$, and we provide an interval for starting values u_0 giving oscillatory solutions (see Theorems 2.3, 2.10, and 2.16). We specify a behaviour of oscillatory solutions in more details (decreasing amplitudes—see Theorems 2.10 and 2.16), and we show conditions which guarantee that oscillatory solutions converge to 0 (Theorem 3.1).

The paper is organized in this manner: Section 2 contains results about existence, uniqueness, and other basic properties of solutions of the problem (1.7), (1.13). These results which mainly concern damped solutions are taken from [18] and extended or modified a little. We also provide here new conditions for the existence of oscillatory solutions in Theorem 2.16. Section 3 is devoted to asymptotic properties of oscillatory solutions, and the main result is contained in Theorem 3.1.

2. Solutions of the Initial Problem (1.7), (1.13)

Let us give an account of this section in more details. The main objective of this paper is to characterize asymptotic properties of oscillatory solutions of the problem (1.7), (1.13). In order to present more complete results about the solutions, we start this section with the unique solvability of the problem (1.7), (1.13) on $[0,\infty)$ (Theorem 2.1). Having such global solutions, we have proved (see papers [15–18]) that oscillatory solutions of the problem (1.7), (1.13) can be found just in the class of damped solutions of this problem. Therefore, we give here one result about the existence of damped solutions (Theorem 2.3). Example 2.5 shows that there are damped solutions which are not oscillatory. Consequently, we bring results about the existence of oscillatory solutions in the class of damped solutions. This can be found in Theorem 2.10, which is an extension of Theorem 3.4 of [18] and in Theorem 2.16, which are new. Theorems 2.10 and 2.16 cover different classes of equations which is illustrated by examples.

to Contents, page 1

Theorem 2.1 (existence and uniqueness). *Assume that* (1.2)–(1.5), (1.8), (1.9) *hold and that there exists* $C_L \in (0, \infty)$ *such that*

$$0 \leq f(x) \leq C_L \quad \text{for } x \geq L \tag{2.1}$$

then the initial problem (1.7), (1.13) *has a unique solution* u*. The solution* u *satisfies*

$$\begin{aligned} u(t) &\geq u_0 \quad \text{if } u_0 < 0, \\ u(t) &> \overline{B} \quad \text{if } u_0 \geq 0, \end{aligned} \quad \text{for } t \in [0, \infty). \tag{2.2}$$

Proof. Let $u_0 < 0$, then the assertion is contained in Theorem 2.1 of [18]. Now, assume that $u_0 \in [0, L]$, then the proof of Theorem 2.1 in [18] can be slightly modified. □

For close existence results, see also Chapters 13 and 14 of [26], where this kind of equations is studied.

Remark 2.2. Clearly, for $u_0 = 0$ and $u_0 = L$, the problem (1.7), (1.13) has a unique solution $u \equiv 0$ and $u \equiv L$, respectively. Since $f \in \mathrm{Lip}_{\mathrm{loc}}(\mathbb{R})$, no solution of the problem (1.7), (1.13) with $u_0 < 0$ or $u_0 \in (0, L)$ can touch the constant solutions $u \equiv 0$ and $u \equiv L$.

In particular, assume that $C \in \{0, L\}$, $a > 0$, u is a solution of the problem (1.7), (1.13) with $u_0 < L$, $u_0 \neq 0$, and (1.2), (1.8), and (1.9) hold. If $u(a) = C$, then $u'(a) \neq 0$, and if $u'(a) = 0$, then $u(a) \neq C$.

The next theorem provides an extension of Theorem 2.4 in [18].

Theorem 2.3 (existence of damped solutions). *Assume that* (1.2)–(1.5), (1.8), *and* (1.9) *hold, then for each* $u_0 \in [\overline{B}, L)$, *the problem* (1.7), (1.13) *has a unique solution. This solution is damped.*

Proof. First, assume that there exists $C_L > 0$ such that f satisfies (2.1), then, by Theorem 2.1, the problem (1.7), (1.13) has a unique solution u satisfying (2.2). Assume that u is not damped, that is,

$$\sup\{u(t) : t \in [0, \infty)\} \geq L. \tag{2.3}$$

By (1.3)–(1.5), the inequality $F(u_0) \leq F(L)$ holds. Since u fulfils (1.7), we have

$$u''(t) + \frac{p'(t)}{p(t)} u'(t) = f(u(t)) \quad \text{for } t \in (0, \infty). \tag{2.4}$$

Multiplying (2.4) by u' and integrating between 0 and $t > 0$, we get

$$0 < \frac{u'^2(t)}{2} + \int_0^t \frac{p'(s)}{p(s)} u'^2(s) \mathrm{d}s = F(u_0) - F(u(t)), \quad t \in (0, \infty), \tag{2.5}$$

to Contents, page 1

and consequently

$$0 < \int_0^t \frac{p'(s)}{p(s)} u'^2(s)\mathrm{d}s \le F(u_0) - F(u(t)), \quad t \in (0,\infty). \tag{2.6}$$

By (2.3), we can find that $b \in (0,\infty]$ such that $u(b) \ge L$, $(u(\infty) = \limsup_{t\to\infty} u(t))$, and hence, according to (1.5),

$$0 < \int_0^b \frac{p'(s)}{p(s)} u'^2(s)\mathrm{d}s \le F(u_0) - F(u(b)) \le F(B) - F(L) \le 0, \tag{2.7}$$

which is a contradiction. We have proved that $\sup\{u(t) : t \in [0,\infty)\} < L$, that is, u is damped. Consequently, assumption (2.1) can be omitted. □

Example 2.4. Consider the equation

$$u'' + \frac{2}{t}u' = u(u-1)(u+2), \tag{2.8}$$

which is relevant to applications in [9–11]. Here, $p(t) = t^2$, $f(x) = x(x-1)(x+2)$, $L_0 = -2$, and $L = 1$. Hence $f(\mathrm{x}) < 0$ for $x \in (0,1)$, $f(x) > 0$ for $x \in (-2,0)$, and

$$F(x) = -\int_0^x f(z)\mathrm{d}z = -\frac{x^4}{4} - \frac{x^3}{3} + x^2. \tag{2.9}$$

Consequently, F is decreasing and positive on $[-2,0)$ and increasing and positive on $(0,1]$. Since $F(1) = 5/12$ and $F(-1) = 13/12$, there exists a unique $\overline{B} \in (-1,0)$ such that $F(\overline{B}) = 5/12 = F(1)$. We can see that all assumptions of Theorem 2.3 are fulfilled and so, for each $u_0 \in [\overline{B},1)$, the problem (2.8), (1.13) has a unique solution which is damped. We will show later (see Example 2.11), that each damped solution of the problem (2.8), (1.13) is oscillatory.

In the next example, we will show that damped solutions can be nonzero and monotonous on $[0,\infty)$ with a limit equal to zero at ∞. Clearly, such solutions are not oscillatory.

Example 2.5. Consider the equation

$$u'' + \frac{3}{t}u' = f(u), \tag{2.10}$$

where

$$f(x) = \begin{cases} -x^3 & \text{for } x \le 1, \\ x-2 & \text{for } x \in (1,3), \\ 1 & \text{for } x \ge 3. \end{cases} \tag{2.11}$$

to Contents, page 1

We see that $p(t) = t^3$ in (2.10) and the functions f and p satisfy conditions (1.2)–(1.5), (1.8), and (1.9) with $L = 2$. Clearly, $L_0 = -\infty$. Further,

$$F(x) = -\int_0^x f(z)\mathrm{d}z = \begin{cases} \dfrac{x^4}{4} & \text{for } x \le 1, \\ -\dfrac{x^2}{2} + 2x - \dfrac{5}{4} & \text{for } x \in (1,3), \\ -x + \dfrac{13}{4} & \text{for } x \ge 3. \end{cases} \tag{2.12}$$

Since $F(L) = F(2) = 3/4$, assumption (1.5) yields $F(\overline{B}) = \overline{B}^4/4 = 3/4$ and $\overline{B} = -3^{1/4}$. By Theorem 2.3, for each $u_0 \in [-3^{1/4}, 2)$, the problem (2.10), (1.13) has a unique solution u which is damped. On the other hand, we can check by a direct computation that for each $u_0 \le 1$ the function

$$u(t) = \frac{8u_0}{8 + u_0^2 t^2}, \quad t \in [0, \infty) \tag{2.13}$$

is a solution of equation (2.10) and satifies conditions (1.13). If $u_0 < 0$, then $u < 0$, $u' > 0$ on $(0, \infty)$, and if $u_0 \in (0, 1]$, then $u > 0$, $u' < 0$ on $(0, \infty)$. In both cases, $\lim_{t \to \infty} u(t) = 0$.

In Example 2.5, we also demonstrate that there are equations fulfilling Theorem 2.3 for which all solutions with $u_0 < L$, not only those with $u_0 \in [\overline{B}, L)$, are damped. Some additional conditions giving, moreover, bounding homoclinic solutions and escape solutions are presented in [15–17].

In our further investigation of asymptotic properties of damped solutions the following lemmas are useful.

Lemma 2.6. *Assume* (1.2), (1.8), *and* (1.9). *Let u be a damped solution of the problem* (1.7), (1.13) *with $u_0 \in (L_0, L)$ which is eventually positive or eventually negative, then*

$$\lim_{t \to \infty} u(t) = 0, \qquad \lim_{t \to \infty} u'(t) = 0. \tag{2.14}$$

Proof. Let u be eventually positive, that is, there exists $t_0 \ge 0$ such that

$$u(t) > 0 \quad \text{for } t \in [t_0, \infty). \tag{2.15}$$

Denote $\theta = \inf\{t_0 \ge 0 : u(t) > 0,\ t \in [t_0, \infty)\}$.

Let $\theta > 0$, then $u(\theta) = 0$ and, by Remark 2.2, $u'(\theta) > 0$. Assume that $u' > 0$ on (θ, ∞), then u is increasing on (θ, ∞), and there exists $\lim_{t \to \infty} u(t) = \ell \in (0, L)$. Multiplying (2.4) by u', integrating between θ and t, and using notation (1.4), we obtain

$$\frac{u'^2(t)}{2} + \int_\theta^t \frac{p'(s)}{p(s)} u'^2(s)\mathrm{d}s = F(u_0) - F(u(t)), \quad t \in (\theta, \infty). \tag{2.16}$$

to Contents, page 1

Letting $t \to \infty$, we get

$$\lim_{t\to\infty} \frac{u'^2(t)}{2} = -\lim_{t\to\infty} \int_\theta^t \frac{p'(s)}{p(s)} u'^2(s)\mathrm{d}s + F(u_0) - F(\ell). \tag{2.17}$$

Since the function $\int_\theta^t (p'(s)/p(s))u'^2(s)\mathrm{d}s$ is positive and increasing, it follows that it has a limit at ∞, and hence there exists also $\lim_{t\to\infty} u'(t) \geq 0$. If $\lim_{t\to\infty} u'(t) > 0$, then $L > l = \lim_{t\to\infty} u(t) = \infty$, which is a contradiction. Consequently

$$\lim_{t\to\infty} u'(t) = 0. \tag{2.18}$$

Letting $t \to \infty$ in (2.4) and using (1.2), (1.9) and $\ell \in (0, L)$, we get $\lim_{t\to\infty} u''(t) = f(\ell) < 0$, and so $\lim_{t\to\infty} u'(t) = -\infty$, which is contrary to (2.18). This contradiction implies that the inequality $u' > 0$ on (θ, ∞) cannot be satisfied and that there exists $a > \theta$ such that $u'(a) = 0$. Since $u > 0$ on (a, ∞), we get by (1.2), (1.7), and (1.13) that $(pu')' < 0$ on (a, ∞). Due to $p(a)u'(a) = 0$, we see that $u' < 0$ on (a, ∞). Therefore, u is decreasing on (a, ∞) and $\lim_{t\to\infty} u(t) = \ell_0 \in [0, L)$. Using (2.16) with a in place of θ, we deduce as above that (2.18) holds and that $\lim_{t\to\infty} u''(t) = f(\ell_0) = 0$. Consequently, $\ell_0 = 0$. We have proved that (2.14) holds provided $\theta > 0$.

If $\theta = 0$, then we take $a = 0$ and use the above arguments. If u is eventually negative, we argue similarly. □

Lemma 2.7. *Assume* (1.2)–(1.5), (1.8), (1.9), *and*

$$p \in C^2(0, \infty), \quad \limsup_{t\to\infty} \left| \frac{p''(t)}{p'(t)} \right| < \infty, \tag{2.19}$$

$$\lim_{x\to 0+} \frac{f(x)}{x} < 0. \tag{2.20}$$

Let u be a solution of the problem (1.7), (1.13) *with $u_0 \in (0, L)$, then there exists $\delta_1 > 0$ such that*

$$u(\delta_1) = 0, \qquad u'(t) < 0 \quad \textit{for } t \in (0, \delta_1]. \tag{2.21}$$

Proof. Assume that such δ_1 does not exist, then u is positive on $[0, \infty)$ and, by Lemma 2.6, u satisfies (2.14). We define a function

$$v(t) = \sqrt{p(t)}u(t), \quad t \in [0, \infty). \tag{2.22}$$

By (2.19), we have $v \in C^2(0, \infty)$ and

$$v'(t) = \frac{p'(t)u(t)}{2\sqrt{p(t)}} + \sqrt{p(t)}u'(t), \tag{2.23}$$

$$v''(t) = v(t)\left[\frac{1}{2}\frac{p''(t)}{p(t)} - \frac{1}{4}\left(\frac{p'(t)}{p(t)}\right)^2 + \frac{f(u(t))}{u(t)}\right], \quad t \in (0, \infty). \tag{2.24}$$

to Contents, page 1

By (1.9) and (2.19), we get

$$\lim_{t\to\infty}\left[\frac{1}{2}\frac{p''(t)}{p(t)}-\frac{1}{4}\left(\frac{p'(t)}{p(t)}\right)^2\right]=\frac{1}{2}\lim_{t\to\infty}\frac{p''(t)}{p'(t)}\cdot\frac{p'(t)}{p(t)}=0. \tag{2.25}$$

Since u is positive on $(0,\infty)$, conditions (2.14) and (2.20) yield

$$\lim_{t\to\infty}\frac{f(u(t))}{u(t)}=\lim_{x\to 0+}\frac{f(x)}{x}<0. \tag{2.26}$$

Consequently, there exist $\omega>0$ and $R>0$ such that

$$\frac{1}{2}\frac{p''(t)}{p(t)}-\frac{1}{4}\left(\frac{p'(t)}{p(t)}\right)^2+\frac{f(u(t))}{u(t)}<-\omega \quad \text{for } t\geq R. \tag{2.27}$$

By (2.22), v is positive on $(0,\infty)$ and, due to (2.24) and (2.27), we get

$$v''(t)<-\omega v(t)<0 \quad \text{for } t\geq R. \tag{2.28}$$

Thus, v' is decreasing on $[R,\infty)$ and $\lim_{t\to\infty}v'(t)=V$. If $V<0$, then $\lim_{t\to\infty}v(t)=-\infty$, contrary to the positivity of v. If $V\geq 0$, then $v'>0$ on $[R,\infty)$ and $v(t)\geq v(R)>0$ for $t\in[R,\infty)$. Then (2.28) yields $0>-\omega v(R)\geq-\omega v(t)>v''(t)$ for $t\in[R,\infty)$. We get $\lim_{t\to\infty}v'(t)=-\infty$ which contradicts $V\geq 0$. The obtained contradictions imply that u has at least one zero in $(0,\infty)$. Let $\delta_1>0$ be the first zero of u. Then $u>0$ on $[0,\delta_1)$ and, by (1.2) and (1.7), $u'<0$ on $(0,\delta_1)$. Due to Remark 2.2, we have also $u'(\delta_1)<0$. □

For negative starting value, we can prove a dual lemma by similar arguments.

Lemma 2.8. *Assume* (1.2)–(1.5), (1.8), (1.9), (2.19) *and*

$$\lim_{x\to 0-}\frac{f(x)}{x}<0. \tag{2.29}$$

Let u be a solution of the problem (1.7), (1.13) *with $u_0\in(L_0,0)$, then there exists $\theta_1>0$ such that*

$$u(\theta_1)=0, \qquad u'(t)>0 \quad \textit{for } t\in(0,\theta_1]. \tag{2.30}$$

The arguments of the proof of Lemma 2.8 can be also found in the proof of Lemma 3.1 in [18], where both (2.20) and (2.29) were assumed. If one argues as in the proofs of Lemmas 2.7 and 2.8 working with a_1, A_1 and b_1, B_1 in place of 0, and u_0, one gets the next corollary.

to Contents, page 1

Corollary 2.9. *Assume* (1.2)–(1.5), (1.8), (1.9), (2.19), (2.20), *and* (2.29). *Let u be a solution of the problem* (1.7), (1.13) *with $u_0 \in (L_0, 0) \cup (0, L)$.*

(I) Assume that there exist $b_1 > 0$ and $B_1 \in (L_0, 0)$ such that

$$u(b_1) = B_1, \quad u'(b_1) = 0, \tag{2.31}$$

then there exists $\theta > b_1$ such that

$$u(\theta) = 0, \qquad u'(t) > 0 \quad \textit{for } t \in (b_1, \theta]. \tag{2.32}$$

(II) Assume that there exist $a_1 > 0$ and $A_1 \in (0, L)$ such that

$$u(a_1) = A_1, \quad u'(a_1) = 0, \tag{2.33}$$

then there exists $\delta > a_1$ such that

$$u(\delta) = 0, \qquad u'(t) < 0 \quad \textit{for } t \in (a_1, \delta]. \tag{2.34}$$

Note that if all conditions of Lemmas 2.7 and 2.8 are satisfied, then each solution of the problem (1.7), (1.13) with $u_0 \in (L_0, 0) \cup (0, L)$ has at least one simple zero in $(0, \infty)$. Corollary 2.9 makes possible to construct an unbounded sequence of all zeros of any damped solution u. In addition, these zeros are simple (see the proof of Theorem 2.10). In such a case, u has either a positive maximum or a negative minimum between each two neighbouring zeros. If we denote sequences of these maxima and minima by $\{A_n\}_{n=1}^{\infty}$ and $\{B_n\}_{n=1}^{\infty}$, respectively, then we call the numbers $|A_n - B_n|$, $n \in \mathbb{N}$ amplitudes of u.

In [18], we give conditions implying that each damped solution of the problem (1.7), (1.13) with $u_0 < 0$ has an unbounded set of zeros and decreasing sequence of amplitudes. Here, there is an extension of this result for $u_0 \in (0, L)$.

Theorem 2.10 (existence of oscillatory solutions I). *Assume that* (1.2)–(1.5), (1.8), (1.9), (2.19), (2.20), *and* (2.29) *hold, Then each damped solution of the problem* (1.7), (1.13) *with $u_0 \in (L_0, 0) \cup (0, L)$ is oscillatory and its amplitudes are decreasing.*

Proof. For $u_0 < 0$, the assertion is contained in Theorem 3.4 of [18]. Let u be a damped solution of the problem (1.7), (1.13) with $u_0 \in (0, L)$. By (2.2) and Definition 1.2, we can find $L_1 \in (0, L)$ such that

$$\overline{B} < u(t) \leq L_1 \quad \text{for } t \in [0, \infty). \tag{2.35}$$

Step 1. Lemma 2.7 yields $\delta_1 > 0$ satisfying (2.21). Hence, there exists a maximal interval (δ_1, b_1) such that $u' < 0$ on (δ_1, b_1). If $b_1 = \infty$, then u is eventually negative and decreasing. On the other hand, by Lemma 2.6, u satisfies (2.14). But this is not possible. Therefore, $b_1 < \infty$ and there exists $B_1 \in (\overline{B}, 0)$ such that (2.31) holds. Corollary 2.9 yields $\theta_1 > b_1$ satisfying (2.32) with $\theta = \theta_1$. Therefore, u has just one negative local minimum $B_1 = u(b_1)$ between its first zero δ_1 and second zero θ_1.

to Contents, page 1

Step 2. By (2.32) there exists a maximal interval (θ_1, a_1), where $u' > 0$. If $a_1 = \infty$, then u is eventually positive and increasing. On the other hand, by Lemma 2.6, u satisfies (2.14). We get a contradiction. Therefore $a_1 < \infty$ and there exists $A_1 \in (0, L)$ such that (2.33) holds. Corollary 2.9 yields $\delta_2 > a_1$ satisfying (2.34) with $\delta = \delta_2$. Therefore u has just one positive maximum $A_1 = u(a_1)$ between its second zero θ_1 and third zero δ_2.

Step 3. We can continue as in Steps 1 and 2 and get the sequences $\{A_n\}_{n=1}^{\infty} \subset (0, L)$ and $\{B_n\}_{n=1}^{\infty} \subset [u_0, 0)$ of positive local maxima and negative local minima of u, respectively. Therefore u is oscillatory. Using arguments of the proof of Theorem 3.4 of [18], we get that the sequence $\{A_n\}_{n=1}^{\infty}$ is decreasing and the sequence $\{B_n\}_{n=1}^{\infty}$ is increasing. In particular, we use (2.5) and define a Lyapunov function V_u by

$$V_u(t) = \frac{u'^2(t)}{2} + F(u(t)) = F(u_0) - \int_0^t \frac{p'(s)}{p(s)} u'^2(s) \mathrm{d}s, \quad t \in (0, \infty), \tag{2.36}$$

then

$$V_u(t) > 0, \quad V_u'(t) = -\frac{p'(t)}{p(t)} u'^2(t) \le 0 \quad \text{for } t \in (0, \infty), \tag{2.37}$$

$$V_u'(t) < 0 \quad \text{for } t \in (0, \infty),\ t \neq a_n, b_n,\ n \in \mathbb{N}. \tag{2.38}$$

Consequently,

$$c_u := \lim_{t \to \infty} V_u(t) \ge 0. \tag{2.39}$$

So, sequences $\{V_u(a_n)\}_{n=1}^{\infty} = \{F(A_n)\}_{n=1}^{\infty}$ and $\{V_u(b_n)\}_{n=1}^{\infty} = \{F(B_n)\}_{n=1}^{\infty}$ are decreasing and

$$\lim_{n \to \infty} F(A_n) = \lim_{n \to \infty} F(B_n) = c_u. \tag{2.40}$$

Finally, due to (1.4), the sequence $\{A_n\}_{n=1}^{\infty}$ is decreasing and the sequence $\{B_n\}_{n=1}^{\infty}$ is increasing. Hence, the sequence of amplitudes $\{A_n - B_n\}_{n=1}^{\infty}$ is decreasing, as well. □

Example 2.11. Consider the problem (1.7), (1.13), where $p(t) = t^2$ and $f(x) = x(x-1)(x+2)$. In Example 2.4, we have shown that (1.2)–(1.5), (1.8), and (1.9) with $L_0 = -2$, $L = 1$ are valid. Since

$$\begin{gathered} \lim_{t \to \infty} \frac{p''(t)}{p'(t)} = \lim_{t \to \infty} \frac{1}{t} = 0, \\ \lim_{x \to 0} \frac{f(x)}{x} = \lim_{x \to 0} (x-1)(x+2) = -2 < 0, \end{gathered} \tag{2.41}$$

to Contents, page 1

we see that (2.19), (2.20), and (2.29) are satisfied. Therefore, by Theorem 2.10, each damped solution of (2.8), (1.13) with $u_0 \in (-2,0) \cup (0,1)$ is oscillatory and its amplitudes are decreasing.

Example 2.12. Consider the problem (1.7), (1.13), where

$$p(t) = \frac{t^k}{1+t^\ell}, \quad k > \ell \geq 0,$$
$$f(x) = \begin{cases} x(x-1)(x+3), & \text{for } x \leq 0, \\ x(x-1)(x+4), & \text{for } x > 0, \end{cases} \tag{2.42}$$

then $L_0 = -3$, $L = 1$,

$$\lim_{t\to\infty} \frac{p''(t)}{p'(t)} = 0, \qquad \lim_{x\to 0-} \frac{f(x)}{x} = -3, \qquad \lim_{x\to 0+} \frac{f(x)}{x} = -4. \tag{2.43}$$

We can check that also all remaining assumptions of Theorem 2.10 are satisfied, and this theorem is applicable here.

Assume that f does not fulfil (2.20) and (2.29). It occurs, for example, if $f(x) = -|x|^\alpha \operatorname{sign} x$ with $\alpha > 1$ for x in some neighbourhood of 0, then Theorem 2.10 cannot be applied. Now, we will give another sufficient conditions for the existence of oscillatory solutions. For this purpose, we introduce the following lemmas.

Lemma 2.13. *Assume* (1.2)–(1.5), (1.8), (1.9), *and*

$$\int_1^\infty \frac{1}{p(s)} \mathrm{d}s = \infty, \tag{2.44}$$

$$\exists \epsilon > 0 : f \in C^1(0,\epsilon), \qquad f' \leq 0 \quad \text{on } (0,\epsilon). \tag{2.45}$$

Let u be a solution of the problem (1.7), (1.13) *with $u_0 \in (0, L)$, then there exists $\delta_1 > 0$ such that*

$$u(\delta_1) = 0, \qquad u'(t) < 0 \quad \textit{for } t \in (0, \delta_1]. \tag{2.46}$$

Proof. Assume that such δ_1 does not exist, then u is positive on $[0,\infty)$ and, by Lemma 2.6, u satisfies (2.14). In view of (1.7) and (1.2), we have $u' < 0$ on $(0,\infty)$. From (2.45), it follows that there exists $t_0 > 0$ such that

$$0 < u(t) < \epsilon, \quad \text{for } t \in [t_0, \infty). \tag{2.47}$$

to Contents, page 1

Motivated by arguments of [27], we divide (1.7) by $f(u)$ and integrate it over interval $[t_0, t]$. We get

$$\int_{t_0}^{t} \frac{(p(s)u'(s))'}{f(u(s))}\mathrm{d}s = \int_{t_0}^{t} p(s)\mathrm{d}s \quad \text{for } t \in [t_0, \infty). \tag{2.48}$$

Using the per partes integration, we obtain

$$\frac{p(t)u'(t)}{f(u(t))} + \int_{t_0}^{t} \frac{p(s)f'(u(s))u'^2(s)}{f^2(u(s))}\mathrm{d}s = \frac{p(t_0)u'(t_0)}{f(u(t_0))} + \int_{t_0}^{t} p(s)\mathrm{d}s, \quad t \in [t_0, \infty). \tag{2.49}$$

From (1.8) and (1.9), it follows that there exists $t_1 \in (t_0, \infty)$ such that

$$\frac{p(t_0)u'(t_0)}{f(u(t_0))} + \int_{t_0}^{t} p(s)\mathrm{d}s \geq 1, \quad t \in [t_1, \infty), \tag{2.50}$$

and therefore

$$\frac{p(t)u'(t)}{f(u(t))} + \int_{t_0}^{t} \frac{p(s)f'(u(s))u'^2(s)}{f^2(u(s))}\mathrm{d}s \geq 1, \quad t \in [t_1, \infty). \tag{2.51}$$

From the fact that $f'(u(s)) \leq 0$ for $s > t_0$ (see (2.45)), we have

$$\frac{p(t)u'(t)}{f(u(t))} + \int_{t_1}^{t} \frac{p(s)f'(u(s))u'^2(s)}{f^2(u(s))}\mathrm{d}s \geq 1, \quad t \in [t_1, \infty), \tag{2.52}$$

then

$$\frac{p(t)u'(t)}{f(u(t))} \geq 1 - \int_{t_1}^{t} \frac{p(s)f'(u(s))u'^2(s)}{f^2(u(s))}\mathrm{d}s > 0, \quad t \in [t_1, \infty), \tag{2.53}$$

$$\frac{p(t)u'(t)}{f(u(t))\left(1 - \int_{t_1}^{t} p(s)f'(u(s))u'^2(s)f^{-2}(u(s))\mathrm{d}s\right)} \geq 1, \quad t \in [t_1, \infty). \tag{2.54}$$

Multiplying this inequality by $-f'(u(t))u'(t)/f(u(t)) \geq 0$, we get

$$\left(\ln\left(1 - \int_{t_1}^{t} \frac{p(s)f'(u(s))u'^2(s)}{f^2(u(s))}\mathrm{d}s\right)\right)' \geq -\left(\ln\left|f(u(t))\right|\right)', \quad t \in [t_1, \infty), \tag{2.55}$$

and integrating it over $[t_1, t]$, we obtain

$$\ln\left(1 - \int_{t_1}^{t} \frac{p(s)f'(u(s))u'^2(s)}{f^2(u(s))}\mathrm{d}s\right) \geq \ln\left(\frac{f(u(t_1))}{f(u(t))}\right), \tag{2.56}$$

and therefore,

$$1-\int_{t_1}^{t}\frac{p(s)f'(u(s))u'^2(s)}{f^2(u(s))}\mathrm{d}s\geq\frac{f(u(t_1))}{f(u(t))},\quad t\in[t_1,\infty).\tag{2.57}$$

According to (2.53), we have

$$\frac{p(t)u'(t)}{f(u(t))}\geq\frac{f(u(t_1))}{f(u(t))},\quad t\in[t_1,\infty),\tag{2.58}$$

and consequently,

$$u'(t)\leq f(u(t_1))\frac{1}{p(t)},\quad t\in[t_1,\infty).\tag{2.59}$$

Integrating it over $[t_1,t]$, we get

$$u(t)\leq u(t_1)+f(u(t_1))\int_{t_1}^{t}\frac{1}{p(s)}\mathrm{d}s,\quad t\in[t_1,\infty).\tag{2.60}$$

From (2.44), it follows that

$$\lim_{t\to\infty}u(t)=-\infty,\tag{2.61}$$

which is a contradiction. □

By similar arguments, we can prove a dual lemma.

Lemma 2.14. *Assume* (1.2)–(1.5), (1.8), (1.9), (2.44), *and*

$$\exists\epsilon>0: f\in C^1(-\epsilon,0),\quad f'\leq 0\ \text{ on } (-\epsilon,0).\tag{2.62}$$

Let u be a solution of the problem (1.7), (1.13) *with $u_0\in(L_0,0)$, then, there exists $\theta_1>0$ such that*

$$u(\theta_1)=0,\qquad u'(t)>0\quad \textit{for } t\in(0,\theta_1].\tag{2.63}$$

Following ideas before Corollary 2.9, we get the next corollary.

Corollary 2.15. *Assume* (1.2)–(1.5), (1.8), (1.9), (2.44), (2.45), *and* (2.62). *Let u be a solution of the problem* (1.7), (1.13) *with $u_0\in(L_0,0)\cup(0,L)$, then the assertions I and II of Corollary 2.9 are valid.*

Now, we are able to formulate another existence result for oscillatory solutions. Its proof is almost the same as the proof of Theorem 2.10 for $u_0\in(L_0,0)$ and the proof of Theorem 3.4 in [18] for $u_0\in(0,L)$. The only difference is that we use Lemmas 2.13, 2.14, and Corollary 2.15, in place of Lemmas 2.7, 2.8, and Corollary 2.9, respectively.

to Contents, page 1

Theorem 2.16 (existence of oscillatory solutions II). *Assume that* (1.2)–(1.5), (1.8), (1.9), (2.44), (2.45), *and* (2.62) *hold, then each damped solution of the problem* (1.7), (1.13) *with* $u_0 \in (L_0, 0) \cup (0, L)$ *is oscillatory and its amplitudes are decreasing.*

Example 2.17. Let us consider (1.7) with

$$p(t) = t^{\alpha}, \quad t \in [0, \infty),$$

$$f(x) = \begin{cases} -|x|^{\lambda}\operatorname{sgn} x, & x \le 1, \\ x - 2, & x \in (1, 3), \\ 1, & x \ge 3, \end{cases} \tag{2.64}$$

where λ and α are real parameters.

Case 1. Let $\lambda \in (1, \infty)$ and $\alpha \in (0, 1]$, then all assumptions of Theorem 2.16 are satisfied. Note that f satisfies neither (2.20) nor (2.29) and hence Theorem 2.10 cannot be applied.

Case 2. Let $\lambda = 1$ and $\alpha \in (0, \infty)$, then all assumptions of Theorem 2.10 are satisfied. If $\alpha \in (0, 1]$, then also all assumptions of Theorem 2.16 are fulfilled, but for $\alpha \in (1, \infty)$, the function p does not satisfy (2.44), and hence Theorem 2.16 cannot be applied.

3. Asymptotic Properties of Oscillatory Solutions

In Lemma 2.6 we show that if u is a damped solution of the problem (1.7), (1.13) which is not oscillatory then u converges to 0 for $t \to \infty$. In this section, we give conditions under which also oscillatory solutions converge to 0.

Theorem 3.1. *Assume that* (1.2)–(1.5), (1.8), *and* (1.9) *hold and that there exists* $k_0 > 0$ *such that*

$$\liminf_{t \to \infty} \frac{p(t)}{t^{k_0}} > 0, \tag{3.1}$$

then each damped oscillatory solution u *of the problem* (1.7), (1.13) *with* $u_0 \in (L_0, 0) \cup (0, L)$ *satisfies*

$$\lim_{t \to \infty} u(t) = 0, \qquad \lim_{t \to \infty} u'(t) = 0. \tag{3.2}$$

Proof. Consider an oscillatory solution u of the problem (1.7), (1.13) with $u_0 \in (0, L)$.

Step 1. Using the notation and some arguments of the proof of Theorem 2.10, we have the unbounded sequences $\{a_n\}_{n=1}^{\infty}$, $\{b_n\}_{n=1}^{\infty}$, $\{\theta_n\}_{n=1}^{\infty}$, and $\{\delta_n\}_{n=1}^{\infty}$, such that

$$0 < \delta_1 < b_1 < \theta_1 < a_1 < \delta_2 < \cdots < \delta_n < b_n < \theta_n < a_n < \delta_{n+1} < \cdots, \tag{3.3}$$

to Contents, page 1

where $u(\theta_n) = u(\delta_n) = 0$, $u(a_n) = A_n > 0$ is a unique local maximum of u in (θ_n, δ_{n+1}), $u(b_n) = B_n < 0$ is a unique local minimum of u in (δ_n, θ_n), $n \in \mathbb{N}$. Let V_u be given by (2.36) and then (2.39) and (2.40) hold and, by (1.2)–(1.4), we see that

$$\lim_{t\to\infty} u(t) = 0 \Longleftrightarrow c_u = 0. \tag{3.4}$$

Assume that (3.2) does not hold. Then $c_u > 0$. Motivated by arguments of [28], we derive a contradiction in the following steps.

Step 2 (estimates of u). By (2.36) and (2.39), we have

$$\lim_{n\to\infty} \frac{u'^2(\delta_n)}{2} = \lim_{n\to\infty} \frac{u'^2(\theta_n)}{2} = c_u > 0, \tag{3.5}$$

and the sequences $\{u'^2(\delta_n)\}_{n=1}^{\infty}$ and $\{u'^2(\theta_n)\}_{n=1}^{\infty}$ are decreasing. Consider $n \in \mathbb{N}$. Then $u'^2(\delta_n)/2 > c_u$ and there are α_n, β_n satisfying $a_n < \alpha_n < \delta_n < \beta_n < b_n$ and such that

$$u'^2(\alpha_n) = u'^2(\beta_n) = c_u, \quad u'^2(t) > c_u, \quad t \in (\alpha_n, \beta_n). \tag{3.6}$$

Since $V_u(t) > c_u$ for $t > 0$ (see (2.39)), we get by (2.36) and (3.6) the inequalities $c_u/2 + F(u(\alpha_n)) > c_u$ and $c_u/2 + F(u(\beta_n)) > c_u$, and consequently $F(u(\alpha_n)) > c_u/2$ and $F(u(\beta_n)) > c_u/2$. Therefore, due to (1.4), there exists $\tilde{c} > 0$ such that

$$u(\alpha_n) > \tilde{c}, \quad u(\beta_n) < -\tilde{c}, \quad n \in \mathbb{N}. \tag{3.7}$$

Similarly, we deduce that there are $\tilde{\alpha}_n, \tilde{\beta}_n$ satisfying $b_n < \tilde{\alpha}_n < \theta_n < \tilde{\beta}_n < a_{n+1}$ and such that

$$u(\tilde{\alpha}_n) < -\tilde{c}, \quad u\left(\tilde{\beta}_n\right) > \tilde{c}, \quad n \in \mathbb{N}. \tag{3.8}$$

The behaviour of u and inequalities (3.7) and (3.8) yield

$$|u(t)| > \tilde{c}, \qquad t \in [\beta_n, \tilde{\alpha}_n] \cup \left[\tilde{\beta}_n, \alpha_{n+1}\right], \ n \in \mathbb{N}. \tag{3.9}$$

Step 3 (estimates of $\beta_n - \alpha_n$). We prove that there exist $c_0, c_1 \in (0, \infty)$ such that

$$c_0 < \beta_n - \alpha_n < c_1, \quad n \in \mathbb{N}. \tag{3.10}$$

Assume on the contrary that there exists a subsequence satisfying $\lim_{\ell\to\infty}(\beta_\ell - \alpha_\ell) = 0$. By the mean value theorem and (3.7), there is $\xi_\ell \in (\alpha_\ell, \beta_\ell)$ such that $0 < 2\tilde{c} < u(\alpha_\ell) - u(\beta_\ell) = |u'(\xi_\ell)|(\beta_\ell - \alpha_\ell)$. Since $F(u(t)) \geq 0$ for $t \in [0, \infty)$, we get by (2.16) the inequality

$$\left|u'(t)\right| < \sqrt{2F(u_0)}, \quad t \in [0, \infty), \tag{3.11}$$

to Contents, page 1

and consequently

$$0 < 2\tilde{c} \leq \sqrt{2F(u_0)} \lim_{\ell \to \infty} (\beta_\ell - \alpha_\ell) = 0, \tag{3.12}$$

which is a contradiction. So, c_0 satisfying (3.10) exists. Using the mean value theorem again, we can find $\tau_n \in (\alpha_n, \delta_n)$ such that $u(\delta_n) - u(\alpha_n) = u'(\tau_n)(\delta_n - \alpha_n)$ and, by (3.6),

$$\delta_n - \alpha_n = \frac{-u(\alpha_n)}{u'(\tau_n)} = \frac{u(\alpha_n)}{|u'(\tau_n)|} < \frac{A_1}{\sqrt{c_u}}. \tag{3.13}$$

Similarly, we can find $\eta_n \in (\delta_n, \beta_n)$ such that

$$\beta_n - \delta_n = \frac{u(\beta_n)}{u'(\eta_n)} = \frac{|u(\beta_n)|}{|u'(\eta_n)|} < \frac{|B_1|}{\sqrt{c_u}}. \tag{3.14}$$

If we put $c_1 = (A_1 + |B_1|)/\sqrt{c_u}$, then (3.10) is fulfilled. Similarly, we can prove

$$c_0 < \tilde{\beta}_n - \tilde{\alpha}_n < c_1, \quad n \in \mathbb{N}. \tag{3.15}$$

Step 4 (estimates of $\alpha_{n+1} - \alpha_n$). We prove that there exist $c_2 \in (0, \infty)$ such that

$$\alpha_{n+1} - \alpha_n < c_2, \quad n \in \mathbb{N}. \tag{3.16}$$

Put $m_1 = \min\{f(x) : B_1 \leq x \leq -\tilde{c}\} > 0$. By (3.9), $B_1 \leq u(t) < -\tilde{c}$ for $t \in [\beta_n, \tilde{\alpha}_n]$, $n \in \mathbb{N}$. Therefore,

$$f(u(t)) \geq m_1, \quad t \in [\beta_n, \tilde{\alpha}_n], \ n \in \mathbb{N}. \tag{3.17}$$

Due to (1.9), we can find $t_1 > 0$ such that

$$\frac{p'(t)}{p(t)} \sqrt{2F(u_0)} < \frac{m_1}{2}, \quad t \in [t_1, \infty). \tag{3.18}$$

Let $n_1 \in \mathbb{N}$ fulfil $\alpha_{n_1} \geq t_1$, then, according to (2.4), (3.11), (3.17), and (3.18), we have

$$u''(t) > -\frac{m_1}{2} + m_1 = \frac{m_1}{2}, \quad t \in [\beta_n, \tilde{\alpha}_n], \ n \geq n_1. \tag{3.19}$$

Integrating (3.19) from b_n to β_n and using (3.6), we get $2\sqrt{c_u} > m_1(b_n - \beta_n)$ for $n \geq n_1$. Similarly we get $2\sqrt{c_u} > m_1(\tilde{\alpha}_n - b_n)$ for $n \geq n_1$. Therefore

$$\frac{4}{m_1}\sqrt{c_u} > \tilde{\alpha}_n - \beta_n, \quad n \geq n_1. \tag{3.20}$$

to Contents, page 1

By analogy, we put $m_2 = \min\{-f(x) : \widetilde{c} \le x \le A_1\} > 0$ and prove that there exists $n_2 \in \mathbb{N}$ such that

$$\frac{4}{m_2}\sqrt{c_u} > \alpha_{n+1} - \widetilde{\beta}_n, \quad n \ge n_2. \tag{3.21}$$

Inequalities (3.10), (3.15), (3.20), and (3.21) imply the existence of c_2 fulfilling (3.16).

Step 5 (construction of a contradiction). Choose $t_0 > c_1$ and integrate the equality in (2.37) from t_0 to $t > t_0$. We have

$$V_u(t) = V_u(t_0) - \int_{t_0}^{t} \frac{p'(\tau)}{p(\tau)} u'^2(\tau) d\tau, \qquad t \ge t_0. \tag{3.22}$$

Choose $n_0 \in \mathbb{N}$ such that $\alpha_{n_0} > t_0$. Further, choose $n \in \mathbb{N}$, $n > n_0$ and assume that $t > \beta_n$, then, by (3.6),

$$\begin{aligned}\int_{t_0}^{t} \frac{p'(\tau)}{p(\tau)} u'^2(\tau) d\tau &> \sum_{j=n_0}^{n} \int_{\alpha_j}^{\beta_j} \frac{p'(\tau)}{p(\tau)} u'^2(\tau) d\tau \\ &> c_u \sum_{j=n_0}^{n} \int_{\alpha_j}^{\beta_j} \frac{p'(\tau)}{p(\tau)} d\tau = c_u \sum_{j=n_0}^{n} [\ln p(\tau)]_{\alpha_j}^{\beta_j}.\end{aligned} \tag{3.23}$$

By virtue of (3.1) there exists $c_3 > 0$ such that $p(t)/t^{k_0} > c_3$ for $t \in [t_0, \infty)$. Thus, $\ln p(t) > \ln c_3 + k_0 \ln t$ and

$$\int_{t_0}^{t} \frac{p'(\tau)}{p(\tau)} u'^2(\tau) d\tau > c_u \sum_{j=n_0}^{n} [\ln c_3 + k_0 \ln t]_{\alpha_j}^{\beta_j} = c_u k_0 \sum_{j=n_0}^{n} \ln \frac{\beta_j}{\alpha_j}. \tag{3.24}$$

Due to (3.10) and $c_1 < \alpha_{n_0}$, we have

$$1 < \frac{\beta_j}{\alpha_j} < 1 + \frac{c_1}{\alpha_j} < 2, \quad j = n_0, \ldots, n, \tag{3.25}$$

and the mean value theorem yields $\xi_j \in (1, 2)$ such that

$$\ln \frac{\beta_j}{\alpha_j} = \left(\frac{\beta_j}{\alpha_j} - 1\right)\frac{1}{\xi_j} > \frac{\beta_j - \alpha_j}{2\alpha_j}, \quad j = n_0, \ldots, n. \tag{3.26}$$

By (3.10) and (3.16), we deduce

$$\frac{\beta_j - \alpha_j}{\alpha_j} > \frac{c_0}{\alpha_j}, \quad \alpha_j < jc_2 + \alpha_1, \ j = n_0, \ldots, n. \tag{3.27}$$

to Contents, page 1

Thus,

$$\frac{\beta_j - \alpha_j}{\alpha_j} > \frac{c_0}{jc_2 + \alpha_1}, \quad j = n_0, \dots, n. \tag{3.28}$$

Using (3.24)–(3.28) and letting t to ∞, we obtain

$$\begin{aligned}\int_{t_0}^{\infty} \frac{p'(\tau)}{p(\tau)} u'^2(\tau) \mathrm{d}\tau &\ge c_u k_0 \sum_{n=n_0}^{\infty} \ln \frac{\beta_n}{\alpha_n} \ge \frac{1}{2} c_u k_0 \sum_{n=n_0}^{\infty} \frac{\beta_n - \alpha_n}{\alpha_n} \\ &\ge \frac{1}{2} c_u k_0 \sum_{n=n_0}^{\infty} \frac{c_0}{nc_2 + \alpha_1} = \infty.\end{aligned} \tag{3.29}$$

Using it in (3.22), we get $\lim_{t\to\infty} V_u(t) = -\infty$, which is a contradiction. So, we have proved that $c_u = 0$.

Using (2.4) and (3.4), we have

$$\lim_{t\to\infty} \left(\frac{u'^2(t)}{2} + \int_0^t \frac{p'(s)}{p(s)} u'^2(s) \mathrm{d}s \right) = F(u_0) - F(0) = F(u_0). \tag{3.30}$$

Since the function $\int_0^t (p'(s)/p(s)) u'^2(s) \mathrm{d}s$ is increasing, there exists

$$\lim_{t\to\infty} \int_0^t \frac{p'(s)}{p(s)} u'^2(s) \mathrm{d}s \le F(u_0). \tag{3.31}$$

Therefore, there exists

$$\lim_{t\to\infty} u'^2(t) = \ell^2. \tag{3.32}$$

If $\ell > 0$, then $\lim_{t\to\infty} |u'(t)| = \ell$, which contradicts (3.4). Therefore, $\ell = 0$ and (3.2) is proved.

If $u_0 \in (L_0, 0)$, we argue analogously. □

Acknowledgments

The authors thank the referees for valuable comments and suggestions. This work was supported by the Council of Czech Government MSM 6198959214.

References

[1] V. Bongiorno, L. E. Scriven, and H. T. Davis, "Molecular theory of fluid interfaces," *Journal of Colloid and Interface Science*, vol. 57, pp. 462–475, 1967.

[2] H. Gouin and G. Rotoli, "An analytical approximation of density profile and surface tension of microscopic bubbles for Van Der Waals fluids," *Mechanics Research Communications*, vol. 24, no. 3, pp. 255–260, 1997.

to Contents, page 1

[3] J. D. van der Waals and R. Kohnstamm, *Lehrbuch der Thermodynamik*, vol. 1, Mass and Van Suchtelen, Leipzig, Germany, 1908.

[4] P. C. Fife, *Mathematical Aspects of Reacting and Diffusing Systems*, vol. 28 of *Lecture Notes in Biomathematics*, Springer, Berlin, Germany, 1979.

[5] R. A. Fischer, "The wave of advance of advantegeous genes," *Journal of Eugenics*, vol. 7, pp. 355–369, 1937.

[6] F. F. Abraham, *Homogeneous Nucleation Theory*, Academic Press, New York, NY, USA, 1974.

[7] A. P. Linde, *Particle Physics and in Ationary Cosmology*, Harwood Academic, Chur, Switzerland, 1990.

[8] G. H. Derrick, "Comments on nonlinear wave equations as models for elementary particles," *Journal of Mathematical Physics*, vol. 5, pp. 1252–1254, 1964.

[9] F. Dell'Isola, H. Gouin, and G. Rotoli, "Nucleation of spherical shell-like interfaces by second gradient theory: numerical simulations," *European Journal of Mechanics, B/Fluids*, vol. 15, no. 4, pp. 545–568, 1996.

[10] G. Kitzhofer, O. Koch, P. Lima, and E. Weinmüller, "Efficient numerical solution of the density profile equation in hydrodynamics," *Journal of Scientific Computing*, vol. 32, no. 3, pp. 411–424, 2007.

[11] P. M. Lima, N. B. Konyukhova, A. I. Sukov, and N. V. Chemetov, "Analytical-numerical investigation of bubble-type solutions of nonlinear singular problems," *Journal of Computational and Applied Mathematics*, vol. 189, no. 1-2, pp. 260–273, 2006.

[12] H. Berestycki, P.-L. Lions, and L. A. Peletier, "An ODE approach to the existence of positive solutions for semilinear problems in $\mathbb{R}^N$," *Indiana University Mathematics Journal*, vol. 30, no. 1, pp. 141–157, 1981.

[13] D. Bonheure, J. M. Gomes, and L. Sanchez, "Positive solutions of a second-order singular ordinary differential equation," *Nonlinear Analysis*, vol. 61, no. 8, pp. 1383–1399, 2005.

[14] M. Conti, L. Merizzi, and S. Terracini, "Radial solutions of superlinear equations on $\mathbb{R}^N$. I. A global variational approach," *Archive for Rational Mechanics and Analysis*, vol. 153, no. 4, pp. 291–316, 2000.

[15] I. Rachůnková and J. Tomeček, "Bubble-type solutions of nonlinear singular problems," *Mathematical and Computer Modelling*, vol. 51, no. 5-6, pp. 658–669, 2010.

[16] I. Rachůnková and J. Tomeček, "Strictly increasing solutions of a nonlinear singular differential equation arising in hydrodynamics," *Nonlinear Analysis*, vol. 72, no. 3–4, pp. 2114–2118, 2010.

[17] I. Rachůnková and J. Tomeček, "Homoclinic solutions of singular nonautonomous second-order differential equations," *Boundary Value Problems*, vol. 2009, Article ID 959636, 21 pages, 2009.

[18] I. Rachůnková, J. Tomeček, and J. Stryja, "Oscillatory solutions of singular equations arising in hydrodynamics," *Advances in Difference Equations*, vol. 2010, Article ID 872160, 13 pages, 2010.

[19] J. S. W. Wong, "Second-order nonlinear oscillations: a case history," in *Differential & Difference Equations and Applications*, pp. 1131–1138, Hindawi Publishing Corporation, New York, NY, USA, 2006.

[20] C. H. Ou and J. S. W. Wong, "On existence of oscillatory solutions of second order Emden-Fowler equations," *Journal of Mathematical Analysis and Applications*, vol. 277, no. 2, pp. 670–680, 2003.

[21] P. J. Y. Wong and R. P. Agarwal, "Oscillatory behavior of solutions of certain second order nonlinear differential equations," *Journal of Mathematical Analysis and Applications*, vol. 198, no. 2, pp. 337–354, 1996.

[22] W.-T. Li, "Oscillation of certain second-order nonlinear differential equations," *Journal of Mathematical Analysis and Applications*, vol. 217, no. 1, pp. 1–14, 1998.

[23] M. R. S. Kulenović and Ć. Ljubović, "All solutions of the equilibrium capillary surface equation are oscillatory," *Applied Mathematics Letters*, vol. 13, no. 5, pp. 107–110, 2000.

[24] L. F. Ho, "Asymptotic behavior of radial oscillatory solutions of a quasilinear elliptic equation," *Nonlinear Analysis*, vol. 41, no. 5-6, pp. 573–589, 2000.

[25] M. Bartušek, M. Cecchi, Z. Došlá, and M. Marini, "On oscillatory solutions of quasilinear differential equations," *Journal of Mathematical Analysis and Applications*, vol. 320, no. 1, pp. 108–120, 2006.

[26] D. O'Regan, *Existence Theory for Nonlinear Ordinary differential Equations*, Kluwer Academic, Dodrecht, The Netherlands, 1997.

[27] I. Kiguradze and T. Chanturia, *Asymptotic Properties of Solutions of Nonautonomous Ordinary Differential Equations*, Nauka, Moscow, Russia, 1990.

[28] I. Kiguradze and B. Shekhter, *Singular Boundary Value Problems for Ordinary Differential Equations of the Second Order*, Viniti, Moscow, Russia, 1987.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 605614, 9 pages
doi:10.1155/2011/605614

Research Article

Existence Results for Singular Boundary Value Problem of Nonlinear Fractional Differential Equation

Yujun Cui

Department of Mathematics, Shandong University of Science and Technology, Qingdao 266510, China

Correspondence should be addressed to Yujun Cui, cyj720201@163.com

Received 27 December 2010; Revised 2 March 2011; Accepted 3 March 2011

Academic Editor: Elena Braverman

By applying a fixed point theorem for mappings that are decreasing with respect to a cone, this paper investigates the existence of positive solutions for the nonlinear fractional boundary value problem: $D_{0^+}^{\alpha}u(t) + f(t, u(t)) = 0$, $0 < t < 1$, $u(0) = u'(0) = u'(1) = 0$, where $2 < \alpha < 3$, $D_{0^+}^{\alpha}$ is the Riemann-Liouville fractional derivative.

1. Introduction

Many papers and books on fractional calculus differential equation have appeared recently. Most of them are devoted to the solvability of the linear initial fractional equation in terms of a special function [1–4]. Recently, there has been significant development in the existence of solutions and positive solutions to boundary value problems for fractional differential equations by the use of techniques of nonlinear analysis (fixed point theorems, Leray-Schauder theory, etc.), see [5, 6] and the references therein.

In this paper, we consider the following boundary value problems of the nonlinear fractional differential equation

$$\begin{gathered} D_{0^+}^{\alpha}u(t) + f(t, u(t)) = 0, \quad 0 < t < 1,\ 2 < \alpha < 3, \\ u(0) = u'(0) = u'(1) = 0, \end{gathered} \tag{1.1}$$

where $D_{0^+}^{\alpha}$ is the standard Riemann-Liouville fractional derivative and $f(t, x)$ is singular at

$x = 0$. Our assumptions throughout are

(H_1) $f(t,x) : (0,1) \times (0,\infty) \to [0,\infty)$ is continuous,

(H_2) $f(t,x)$ is decreasing in x, for each fixed t,

(H_3) $\lim_{x\to 0^+} f(t,x) = \infty$ and $\lim_{x\to\infty} f(t,x) = 0$, uniformly on compact subsets of (0,1), and

(H_4) $0 < \int_0^1 f(t, q_\theta(t))dt < \infty$ for all $\theta > 0$ and q_θ as defined in (3.1).

The seminal paper by Gatica et al. [7] in 1989 has had a profound impact on the study of singular boundary value problems for ordinary differential equations (ODEs). They studied singularities of the type in (H_1)–(H_4) for second order Sturm-Louiville problems, and their key result hinged on an application of a particular fixed point theorem for operators which are decreasing with respect to a cone. Various authors have used these techniques to study singular problems of various types. For example, Henderson and Yin [8] as well as Eloe and Henderson [9, 10] have studied right focal, focal, conjugate, and multipoint singular boundary value problems for ODEs. However, as far as we know, no paper is concerned with boundary value problem for fractional differential equation by using this theorem. As a result, the goal of this paper is to fill the gap in this area.

Motivated by the above-mentioned papers and [11], the purpose of this paper is to establish the existence of solutions for the boundary value problem (1.1) by the use of a fixed point theorem used in [7, 11]. The paper has been organized as follows. In Section 2, we give basic definitions and provide some properties of the corresponding Green's function which are needed later. We also state the fixed point theorem from [7] for mappings that are decreasing with respect to a cone. In Section 3, we formulate two lemmas which establish a priori upper and lower bounds on solutions of (1.1). We then state and prove our main existence theorem.

For fractional differential equation and applications, we refer the reader to [1–3]. Concerning boundary value problems (1.1) with ordinary derivative (not fractional one), we refer the reader to [12, 13].

2. Some Preliminaries and a Fixed Point Theorem

For the convenience of the reader, we present here the necessary definitions from fractional calculus theory. These definitions and properties can be found in the literature.

Definition 2.1 (see [3]). The Riemann-Liouville fractional integral of order $\alpha > 0$ of a function $f : (0,\infty) \to R$ is given by

$$I_{0+}^{\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s) ds, \tag{2.1}$$

provided that the right-hand side is pointwise defined on $(0,\infty)$.

Definition 2.2 (see [3]). The Riemann-Liouville fractional derivative of order $\alpha > 0$ of a continuous function $f : (0,\infty) \to R$ is given by

$$D_{0+}^{\alpha} f(t) = \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dt}\right)^n \int_0^t \frac{f(s)}{(t-s)^{\alpha-n+1}} ds, \tag{2.2}$$

to Contents, page 1

where $n-1 \leq \alpha < n$, provided that the right-hand side is pointwise defined on $(0,\infty)$.

Definition 2.3. By a solution of the boundary value problem (1.1) we understand a function $u \in C[0,1]$ such that $D^{\alpha}_{0^+}u$ is continuous on $(0,1)$ and u satisfies (1.1).

Lemma 2.4 (see [3]). *Assume that $u \in C(0,1) \cap L(0,1)$ with a fractional derivative of order $\alpha > 0$ that belongs to $C(0,1) \cap L(0,1)$. Then*

$$I^{\alpha}_{0+}D^{\alpha}_{0+}u(t) = u(t) + c_1 t^{\alpha-1} + c_2 t^{\alpha-2} + \cdots + c_N t^{\alpha-N} \tag{2.3}$$

for some $c_i \in R, i = 1,\ldots,N$, $N = [\alpha]$.

Lemma 2.5. *Given $f \in C[0,1]$, and $2 < \alpha < 3$, the unique solution of*

$$\begin{gathered} D^{\alpha}_{0^+}u(t) + f(t) = 0, \quad 0 < t < 1, \\ u(0) = u'(0) = u'(1) = 0 \end{gathered} \tag{2.4}$$

is

$$u(t) = \int_0^1 G(t,s) f(s) ds, \tag{2.5}$$

where

$$G(t,s) = \begin{cases} \dfrac{t^{\alpha-1}(1-s)^{\alpha-2} - (t-s)^{\alpha-1}}{\Gamma(\alpha)}, & 0 \leq s \leq t \leq 1, \\ \dfrac{t^{\alpha-1}(1-s)^{\alpha-2}}{\Gamma(\alpha)}, & 0 \leq t \leq s \leq 1. \end{cases} \tag{2.6}$$

Proof. We may apply Lemma 2.4 to reduce (2.4) to an equivalent integral equation

$$u(t) = -I^{\alpha}_{0+}f(t) + c_1 t^{\alpha-1} + c_2 t^{\alpha-2} + c_3 t^{\alpha-3} \tag{2.7}$$

for some $c_i \in R,\ i = 1,2,3$. From $u(0) = u'(0) = u'(1) = 0$, one has

$$c_1 = \int_0^1 \frac{(1-s)^{\alpha-2}}{\Gamma(\alpha)} f(s) ds, \quad c_2 = c_3 = 0. \tag{2.8}$$

Therefore, the unique solution of problem (2.4) is

$$\begin{aligned} u(t) &= \int_0^1 \frac{t^{\alpha-1}(1-s)^{\alpha-2}}{\Gamma(\alpha)} f(s) ds - \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s) ds \\ &= \int_0^t \left[\frac{t^{\alpha-1}(1-s)^{\alpha-2} - (t-s)^{\alpha-1}}{\Gamma(\alpha)} \right] f(s) ds + \int_t^1 \frac{t^{\alpha-1}(1-s)^{\alpha-2}}{\Gamma(\alpha)} f(s) ds \\ &= \int_0^1 G(t,s) f(s) ds. \end{aligned} \tag{2.9}$$

□

Lemma 2.6. *The function* $G(t,s)$ *defined by* (2.6) *satisfies the following conditions:*

(i) $G(t,s) > 0$, $0 < t, s < 1$,

(ii) $q(t)G(1,s) \leq G(t,s) \leq G(1,s) = s(1-s)^{\alpha-2}/(\Gamma(\alpha))$ *for* $0 \leq t, s \leq 1$, *where* $q(t) = t^{\alpha-1}$.

Proof. Observing the expression of $G(t,s)$, it is clear that $G(t,s) > 0$ for $0 < t, s < 1$. For given $s \in (0,1)$, $G(t,s)$ is increasing with respect to t. Consequently, $G(t,s) \leq G(1,s)$ for $0 \leq t, s \leq 1$. If $s \leq t$, we have

$$\begin{aligned} G(t,s) &= \frac{t(t-ts)^{\alpha-2} - (t-s)(t-s)^{\alpha-2}}{\Gamma(\alpha)} \\ &\geq \frac{t(t-ts)^{\alpha-2} - (t-s)(t-ts)^{\alpha-2}}{\Gamma(\alpha)} \\ &= \frac{st^{\alpha-2}(1-s)^{\alpha-2}}{\Gamma(\alpha)} \geq \frac{st^{\alpha-1}(1-s)^{\alpha-2}}{\Gamma(\alpha)} = q(t)G(1,s). \end{aligned} \tag{2.10}$$

If $t \leq s$, we have

$$G(t,s) = \frac{t^{\alpha-1}(1-s)^{\alpha-2}}{\Gamma(\alpha)} \geq \frac{st^{\alpha-1}(1-s)^{\alpha-2}}{\Gamma(\alpha)} = q(t)G(1,s). \tag{2.11}$$

□

Let E be a Banach space, $P \subset E$ be a cone in E. Every cone P in E defines a partial ordering in E given by $x \leq y$ if and only if $y - x \in P$. If $x \leq y$ and $x \neq y$, we write $x < y$. A cone P is said to be normal if there exists a constant $N > 0$ such that $\theta \leq x \leq y$ implies $\|x\| \leq N\|y\|$. If P is normal, then every order interval $[x,y] = \{z \in E \mid x \leq z \leq y\}$ is bounded. For the concepts and properties about the cone theory we refer to [14, 15].

Next we state the fixed point theorem due to Gatica et al. [7] which is instrumental in proving our existence results.

Theorem 2.7 (Gatica-Oliker-Waltman fixed point theorem). *Let* E *be a Banach space,* $P \subset E$ *be a normal cone, and* $D \subset P$ *be such that if* $x, y \in D$ *with* $x \leq y$*, then* $[x,y] \subset D$*. Let* $T : D \to P$ *be a continuous, decreasing mapping which is compact on any closed order interval contained in* D*, and suppose there exists an* $x_0 \subset D$ *such that* T^2x_0 *is defined (where* $T^2x_0 = T(Tx_0)$*) and* Tx_0*,* T^2x_0 *are order comparable to* x_0*. Then* T *has a fixed point in* D *provided that either:*

(i) $Tx_0 \leq x_0$ *and* $T^2x_0 \leq x_0$*;*

(ii) $x_0 \leq Tx_0$ *and* $x_0 \leq T^2x_0$*; or*

(iii) *The complete sequence of iterates* $\{T^n x_0\}_{n=0}^{\infty}$ *is defined and there exists* $y_0 \in D$ *such that* $Ty_0 \in D$ *with* $y_0 \leq T^n x_0$ *for all* $n \in \mathbb{N}$.

3. Main Results

In this section, we apply Theorem 2.7 to a sequence of operators that are decreasing with respect to a cone. These obtained fixed points provide a sequence of iterates which converges to a solution of (1.1).

to Contents, page 1

Let the Banach space $E = C[0,1]$ with the maximum norm $\|u\| = \max_{t\in[0,1]}|u(t)|$, and let $P = \{u \in E \mid u(t) \geq 0, t \in [0,1]\}$. P is a norm cone in E. For $\theta > 0$, let

$$q_\theta(t) = \theta \cdot q(t), \tag{3.1}$$

where $q(t)$ is given in Lemma 2.6. Define $D \subset P$ by

$$D = \{u \in P \mid \exists\theta(u) > 0 \text{ such that } u(t) \geq q_\theta(t), t \in [0,1]\}, \tag{3.2}$$

and the integral operator $T : D \to P$ by

$$(Tu)(t) = \int_0^1 G(t,s) f(s,u(s)) ds, \tag{3.3}$$

where $G(t,s)$ is given in (2.6). It suffices to define D as above, since the singularity in f precludes us from defining T on all of P. Furthermore, it can easily be verified that T is well defined. In fact, note that for $u \in D$ there exists $\theta(u) > 0$ such that $u(t) \geq q_\theta(t)$ for all $t \in [0,1]$. Since $f(t,x)$ decreases with respect to x, we see $f(t,u(t)) \leq f(t,q_\theta(t))$ for $t \in [0,1]$. Thus,

$$0 \leq \int_0^1 G(t,s) f(s,u(s)) ds \leq \int_0^1 f(s,q_\theta(s)) ds < \infty. \tag{3.4}$$

Similarly, T is decreasing with respect to D.

Lemma 3.1. *$u \in D$ is a solution of* (1.1) *if and only if $Tu = u$.*

Proof. One direction of the lemma is obviously true. To see the other direction, let $u \in D$. Then $(Tu)(t) = \int_0^1 G(t,s) f(s,u(s)) ds$, and Tu satisfies (1.1). Moreover, by Lemma 2.6, we have

$$\begin{aligned}(Tu)(t) &= \int_0^1 G(t,s) f(s,u(s)) ds \\ &\geq q(t) \int_0^1 G(1,s) f(s,u(s)) ds = q(t)\|Tu\|, \quad \forall t \in [0,1].\end{aligned} \tag{3.5}$$

Thus, there exists some $\theta(Tu)$ such that $(Tu)(t) \geq q_\theta(t)$, which implies that $Tu \in D$. That is, $T : D \to D$.

We now present two lemmas that are required in order to apply Theorem 2.7. The first establishes a priori upper bound on solutions, while the second establishes a priori lower bound on solutions. □

Lemma 3.2. *If f satisfies (H_1)–(H_4), then there exists an $S > 0$ such that $\|u\| \leq S$ for any solution $u \in D$ of* (1.1).

to Contents, page 1

Proof. For the sake of contradiction, suppose that the conclusion is false. Then there exists a sequence $\{u_n\}_{n=1}^{\infty}$ of solutions to (1.1) such that $\|u_n\| \le \|u_{n+1}\|$ with $\lim_{n\to\infty}\|u_n\| = \infty$. Note that for any solution $u_n \in D$ of (1.1), by (3.5), we have

$$u_n(t) = (Tu_n)(t) \ge q(t)\|u_n\|, \quad t \in [0,1],\ n \ge 1. \tag{3.6}$$

Then, assumptions (H_2) and (H_4) yield, for $0 \le t \le 1$ and all $n \ge 1$,

$$\begin{aligned} u_n(t) = (Tu_n)(t) &= \int_0^1 G(t,s) f(s,u_n(s))ds \\ &\le \frac{1}{\Gamma(\alpha)}\int_0^1 s(1-s)^{\alpha-2} f\left(s, q_{\|u_1\|}(s)\right)ds = N, \end{aligned} \tag{3.7}$$

for some $0 < N < +\infty$. In particular, $\|u_n\| \le N$, for all $n \ge 1$, which contradicts $\lim_{n\to\infty}\|u_n\| = \infty$. □

Lemma 3.3. *If f satisfies (H_1)–(H_4), then there exists an $R > 0$ such that $\|u\| \ge R$ for any solution $u \in D$ of* (1.1).

Proof. For the sake of contradiction, suppose $u_n(t) \to 0$ uniformly on $[0,1]$ as $n \to \infty$. Let $M = \inf\{G(t,s) : (t,s) \in [1/4,3/4] \times [1/4,3/4]\} > 0$. From ($H_3$), we see that $\lim_{x\to 0^+} f(t,x) = \infty$ uniformly on compact subsets of $(0,1)$. Hence, there exists some $\delta > 0$ such that for $t \in [1/4,3/4]$ and $0 < x < \delta$, we have $f(t,x) \ge 2/M$. On the other hand, there exists an $n_0 \in N$ such that $n \ge n_0$ implies $0 < u_n(t) < \delta/2$, for $t \in (0,1)$. So, for $t \in [1/4,3/4]$ and $n \ge n_0$,

$$\begin{aligned} u_n(t) = (Tu_n)(t) &= \int_0^1 G(t,s) f(s,u_n(s))ds \ge \int_{1/4}^{3/4} G(t,s) f(s,u_n(s))ds \\ &\ge M\int_{1/4}^{3/4} f\left(s,\frac{\delta}{2}\right)ds \ge M\int_{1/4}^{3/4}\frac{2}{M}ds = 1. \end{aligned} \tag{3.8}$$

But this contradicts the assumption that $\|u_n\| \to 0$ uniformly on $[0,1]$ as $n \to \infty$. Hence, there exists an $R > 0$ such that $R \le \|u\|$. □

We now present the main result of the paper.

Theorem 3.4. *If f satisfies (H_1)–(H_4), then* (1.1) *has at least one positive solution.*

Proof. For each $n \ge 1$, defined $v_n : [0,1] \to [0,+\infty)$ by

$$v_n(t) = \int_0^1 G(t,s) f(s,n)ds. \tag{3.9}$$

By conditions (H_1)–(H_4), for $n \geq 1$,

$$0 < v_{n+1}(t) \leq v_n(t), \quad \text{on } (0,1], \tag{3.10}$$

$$\lim_{n\to\infty} v_n(t) = 0 \quad \text{uniformly on } [0,1]. \tag{3.11}$$

Now define a sequence of functions $f_n : (0,1) \times [0,+\infty)$, $n \geq 1$, by

$$f_n(t,x) = f(t, \max\{x, v_n(t)\}). \tag{3.12}$$

Then, for each $n \geq 1$, f_n is continuous and satisfies (H_2). Furthermore, for $n \geq 1$,

$$\begin{gathered} f_n(t,x) \leq f(t,x) \quad \text{on } (0,1) \times (0,+\infty), \\ f_n(t,x) \leq f(t, v_n(t)) \quad \text{on } (0,1) \times (0,+\infty). \end{gathered} \tag{3.13}$$

Note that f_n has effectively "removed the singularity" in $f(t,x)$ at $x = 0$, then we define a sequence of operators $T_n : P \to P$, $n \geq 1$, by

$$(T_n u)(t) = \int_0^1 G(t,s) f_n(s, u(s)) ds, \ u \in P. \tag{3.14}$$

From standard arguments involving the Arzela-Ascoli Theorem, we know that each T_n is in fact a compact mapping on P. Furthermore, $T_n(0) \geq 0$ and $T_n^2(0) \geq 0$. By Theorem 2.7, for each $n \geq 1$, there exists $u_n \in P$ such that $T_n u_n(x) = u_n(t)$ for $t \in [0,1]$. Hence, for each $n \geq 1$, u_n satisfies the boundary conditions of the problem. In addition, for each u_n,

$$\begin{aligned} (T_n u_n)(t) &= \int_0^1 G(t,s) f_n(s, u_n(s)) ds = \int_0^1 G(t,s) f_n(s, \max\{u_n(s), v_n(s)\}) ds \\ &\leq \int_0^1 G(t,s) f_n(s, v_n(s)) ds \leq T v_n(t), \end{aligned} \tag{3.15}$$

which implies

$$u_n(t) = (T_n u_n)(t) \leq T v_n(t), \quad t \in [0,1], \ n \in \mathrm{N}. \tag{3.16}$$

Arguing as in Lemma 3.2 and using (3.11), it is fairly straightforward to show that there exists an $S > 0$ such that $\|u_n\| \leq S$ for all $n \in \mathbb{N}$. Similarly, we can follow the argument of Lemma 3.3 and (3.5) to show that there exists an $R > 0$ such that

$$u_n(t) \geq q(t)R, \quad \text{on } [0,1], \ \text{for } n \geq 1. \tag{3.17}$$

Since $T : D \to D$ is a compact mapping, there is a subsequence of $\{Tu_n\}$ which converges to some $u^* \in D$. We relabel the subsequence as the original sequence so that $Tu_n \to u^*$ as $n \to \infty$.

to Contents, page 1

To conclude the proof of this theorem, we need to show that

$$\lim_{n\to\infty} \|Tu_n - u_n\| = 0. \tag{3.18}$$

To that end, fixed $\theta = R$, and let $\varepsilon > 0$ be give. By the integrability condition (H_4), there exists $0 < \delta < 1$ such that

$$\int_0^{\delta} s(1-s)^{\alpha-2} f(s, q_{\theta}(s))\,ds < \frac{\Gamma(\alpha)}{2}\varepsilon. \tag{3.19}$$

Further, by (3.11), there exists an n_0 such that, for $n \geq n_0$,

$$v_n(t) \leq q_{\theta}(t) \quad \text{on } [\delta, 1], \tag{3.20}$$

so that

$$v_n(t) \leq q_{\theta}(t) \leq u_n(t) \quad \text{on } [\delta, 1]. \tag{3.21}$$

Thus, for $s \in [\delta, 1]$ and $n \geq n_0$,

$$f_n(s, u_n(s)) = f(s, \max\{u_n(s), v_n(s)\}) = f(s, u_n(s)), \tag{3.22}$$

and for $t \in [0, 1]$,

$$\begin{aligned} Tu_n(t) - u_n(t) &= Tu_n(t) - T_n u_n(t) \\ &= \int_0^1 G(t,s)\left[f(s, u_n(s)) - f_n(s, u_n(s))\right] ds. \end{aligned} \tag{3.23}$$

Thus, for $t \in [0, 1]$,

$$\begin{aligned} |Tu_n(t) - u_n(t)| &\leq \frac{1}{\Gamma(\alpha)}\left[\int_0^{\delta} s(1-s)^{\alpha-2} f(s, u_n(s))ds + \int_0^{\delta} s(1-s)^{\alpha-2} f(s, \max\{u_n(s), v_n(s)\})ds\right] \\ &\leq \frac{1}{\Gamma(\alpha)}\left[\int_0^{\delta} s(1-s)^{\alpha-2} f(s, u_n(s))ds + \int_0^{\delta} s(1-s)^{\alpha-2} f(s, u_n(s))ds\right] \\ &\leq \frac{2}{\Gamma(\alpha)} \int_0^{\delta} s(1-s)^{\alpha-2} f(s, q_{\theta}(s))ds < \varepsilon. \end{aligned} \tag{3.24}$$

Since $t \in [0, 1]$ was arbitrary, we conclude that $\|Tu_n - u_n\| \leq \varepsilon$ for all $n \geq n_0$. Hence, $u^* \in [q_R, S]$ and for $t \in [0, 1]$

$$Tu^*(t) = T\left(\lim_{n\to\infty} Tu_n(t)\right) = T\left(\lim_{n\to\infty} u_n(t)\right) = \lim_{n\to\infty} Tu_n = u^*(t). \tag{3.25}$$

□

to Contents, page 1

Acknowledgments

The authors sincerely thank the reviewers for their valuable suggestions and useful comments that have led to the present improved version of the original paper. The Project Supported by the National Science Foundation of China (10971179) and Research Award Fund for Outstanding Young Scientists of Shandong Province (BS2010SF023).

References

[1] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and Applications of Fractional Differential Equations*, vol. 204 of *North-Holland Mathematics Studies*, Elsevier Science, Amsterdam, The Netherlands, 2006.

[2] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, A Wiley-Interscience Publication, John Wiley & Sons, New York, NY, USA, 1993.

[3] I. Podlubny, *Fractional Differential Equations*, vol. 198 of *Mathematics in Science and Engineering*, Academic Press, San Diego, Calif, USA, 1999.

[4] S. Zhang, "Existence of solution for a boundary value problem of fractional order," *Acta Mathematica Scientia*, vol. 26, no. 2, pp. 220–228, 2006.

[5] Z. Bai and H. Lü, "Positive solutions for boundary value problem of nonlinear fractional differential equation," *Journal of Mathematical Analysis and Applications*, vol. 311, no. 2, pp. 495–505, 2005.

[6] X. Xu, D. Jiang, and C. Yuan, "Multiple positive solutions for the boundary value problem of a nonlinear fractional differential equation," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 10, pp. 4676–4688, 2009.

[7] J. A. Gatica, V. Oliker, and P. Waltman, "Singular nonlinear boundary value problems for second-order ordinary differential equations," *Journal of Differential Equations*, vol. 79, no. 1, pp. 62–78, 1989.

[8] J. Henderson and W. Yin, "Singular $(k, n-k)$ boundary value problems between conjugate and right focal," *Journal of Computational and Applied Mathematics*, vol. 88, no. 1, pp. 57–69, 1998.

[9] P. W. Eloe and J. Henderson, "Singular nonlinear $(k, n-k)$ conjugate boundary value problems," *Journal of Differential Equations*, vol. 133, no. 1, pp. 136–151, 1997.

[10] P. W. Eloe and J. Henderson, "Singular nonlinear multipoint conjugate boundary value problems," *Communications in Applied Analysis*, vol. 2, no. 4, pp. 497–511, 1998.

[11] J. J. DaCunha, J. M. Davis, and P. K. Singh, "Existence results for singular three point boundary value problems on time scales," *Journal of Mathematical Analysis and Applications*, vol. 295, no. 2, pp. 378–391, 2004.

[12] Y. Feng and S. Liu, "Solvability of a third-order two-point boundary value problem," *Applied Mathematics Letters*, vol. 18, no. 9, pp. 1034–1040, 2005.

[13] Q. Yao and Y. Feng, "The existence of solution for a third-order two-point boundary value problem," *Applied Mathematics Letters*, vol. 15, no. 2, pp. 227–232, 2002.

[14] K. Deimling, *Nonlinear Functional Analysis*, Springer, Berlin, Germany, 1985.

[15] D. J. Guo and V. Lakshmikantham, *Nonlinear Problems in Abstract Cones*, vol. 5 of *Notes and Reports in Mathematics in Science and Engineering*, Academic Press, Boston, Mass, USA, 1988.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 182831, 19 pages
doi:10.1155/2011/182831

Research Article

Existence Theory for Pseudo-Symmetric Solution to p-Laplacian Differential Equations Involving Derivative

You-Hui Su,[1] Weili Wu,[1] and Xingjie Yan[2]

[1] *School of Mathematics and Physics, XuZhou Institute of Technology, Xuzhou, Jiangsu 221008, China*
[2] *College of Sciences, China University of Mining and Technology, Xuzhou, Jiangsu 221008, China*

Correspondence should be addressed to Xingjie Yan, yanxj04@cumt.edu.cn

Received 19 November 2010; Accepted 2 May 2011

Academic Editor: Yuri V. Rogovchenko

We all-sidedly consider a three-point boundary value problem for p-Laplacian differential equation with nonlinear term involving derivative. Some new sufficient conditions are obtained for the existence of at least one, triple, or arbitrary odd positive pseudosymmetric solutions by using pseudosymmetric technique and fixed-point theory in cone. As an application, two examples are given to illustrate the main results.

1. Introduction

Recent research results indicate that considerable achievement was made in the existence of positive solutions to dynamic equations; for details, please see [1–6] and the references therein. In particular, the existence of positive pseudosymmetric solutions to p-Laplacian difference and differential equations attract many researchers' attention, such as [7–11]. The reason is that the pseudosymmetry problem not only has theoretical value, such as in the study of metric manifolds [12], but also has practical value itself; for example, we can apply this characteristic into studying the chemistry structure [13]. On another hand, there are much attention paid to the positive solutions of boundary value problems (BVPs) for differential equation with the nonlinear term involved with the derivative explicitly [14–18]. Hence, it is natural to continue study pseudosymmetric solutions to p-Laplacian differential equations with the nonlinear term involved with the first-order derivative explicitly.

First, let us recall some relevant results about BVPs with p-Laplacian, We would like to mention the results of Avery and Henderson [7, 8], Ma and Ge [11] and Sun and Ge [16].

to Contents, page 1

Throughout this paper, we denote the p-Laplacian operator by $\varphi_p(u)$; that is, $\varphi_p(u) = |u|^{p-2}u$ for $p > 1$ with $(\varphi_p)^{-1} = \varphi_q$ and $1/p + 1/q = 1$.

For the three-point BVPs with p-Laplacian

$$\begin{gathered}(\varphi_p(u'(t)))' + h(t)f(t,u(t)) = 0 \quad \text{for } t \in [0,1],\\ u(0) = 0, \quad u(\eta) = u(1),\end{gathered} \tag{1.1}$$

here, $\eta \in (0,1)$ is constant, by using the five functionals fixed point theorem in a cone [19], Avery and Henderson [8] established the existence of at least *three* positive pseudosymmetric solutions to BVPs (1.1). The authors also obtained the similar results in their paper [7] for the discrete case. In addition, Ma and Ge [11] developed the existence of at least *two* positive pseudosymmetric solutions to BVPs (1.1) by using the monotone iterative technique.

For the three-point p-Laplacian BVPs with dependence on the first-order derivative

$$\begin{gathered}(\varphi_p(u'(t)))' + h(t)f(t,u(t),u'(t)) = 0 \quad \text{for } t \in [0,1],\\ u(0) = 0, \quad u(\eta) = u(1),\end{gathered} \tag{1.2}$$

Sun and Ge [16] obtained the existence of at least *two* positive pseudosymmetric solutions to BVPs (1.2) via the monotone iterative technique again. However, it is worth mentioning that the above-mentioned papers [7, 8, 10, 11, 16], the authors only considered results on the existence of positive pseudosymmetric solutions partly, they failed to further provide comprehensive results on the existence of positive pseudosymmetric solutions to p-Laplacian. Naturally, in this paper, we consider the existence of positive pseudosymmetric solutions for p-Laplacian differential equations in all respects.

Motivated by the references [7, 8, 10, 11, 16, 18], in present paper, we consider all-sidedly p-Laplacian BVPs (1.2), using the compression and expansion fixed point theorem [20] and Avery-Peterson fixed point theorem [21]. We obtain that there exist at least *one, triple or arbitrary odd* positive pseudosymmetric solutions to problem (1.2). In particular, we not only get some local properties of pseudosymmetric solutions, but also obtain that the position of pseudosymmetric solutions is determined under some conditions, which is much better than the results in papers [8, 11, 16]. Correspondingly, we generalize and improve the results in papers Avery and Henderson [8]. From the view of applications, two examples are given to illustrate the main results.

Throughout this paper, we assume that

(S1) $f(t,u,u') : [0,1] \times [0,\infty) \times (-\infty,+\infty) \to [0,\infty)$ is continuous, does not vanish identically on interval $[0,1]$, and $f(t,u,u')$ is pseudosymmetric about η on $[0,1]$,

(S2) $h(t) \in L([0,1],[0,\infty))$ is pseudosymmetric about η on $[0,1]$, and does not vanish identically on any closed subinterval of $[0,1]$. Furthermore, $0 < \int_0^1 h(t)dt < \infty$.

2. Preliminaries

In the preceding of this section, we state the definition of cone and several fixed point theorems needed later [20, 22]. In the rest of this section, we will prove that solving BVPs (1.2) is equivalent to finding the fixed points of a completely continuous operator.

to Contents, page 1

We first list the definition of cone and the compression and expansion fixed point theorem [20, 22].

Definition 2.1. Let E be a real Banach space. A nonempty, closed, convex set $P \subset E$ is said to be a cone provided the following conditions are satisfied:

(i) if $x \in P$ and $\lambda \geq 0$, then $\lambda x \in P$,

(ii) if $x \in P$ and $-x \in P$, then $x = 0$.

Lemma 2.2 (see [20, 22]). *Let P be a cone in a Banach space E. Assume that Ω_1, Ω_2 are open bounded subsets of E with $0 \in \Omega_1, \overline{\Omega_1} \subset \Omega_2$. If $A : P \cap (\overline{\Omega_2} \setminus \Omega_1) \to P$ is a completely continuous operator such that either*

(i) $\|Ax\| \leq \|x\|, \forall x \in P \cap \partial\Omega_1$ *and* $\|Ax\| \geq \|x\|, \forall x \in P \cap \partial\Omega_2$, *or*

(ii) $\|Ax\| \geq \|x\|, \forall x \in P \cap \partial\Omega_1$ *and* $\|Ax\| \leq \|x\|, \forall x \in P \cap \partial\Omega_2$.

Then, A has a fixed point in $P \cap (\overline{\Omega_2} \setminus \Omega_1)$.

Given a nonnegative continuous functional γ on a cone P of a real Banach space E, we define, for each $d > 0$, the set $P(\gamma, d) = \{x \in P : \gamma(x) < d\}$.

Let γ and θ be nonnegative continuous convex functionals on P, α a nonnegative continuous concave functional on P, and ψ a nonnegative continuous functional on P respectively. We define the following convex sets:

$$\begin{aligned} P(\gamma, \alpha, b, d) &= \{x \in P : b \leq \alpha(x),\ \gamma(x) \leq d\}, \\ P(\gamma, \theta, \alpha, b, c, d) &= \{x \in P : b \leq \alpha(x),\ \theta(x) \leq c, \gamma(x) \leq d\}, \end{aligned} \tag{2.1}$$

and a closed set $R(\gamma, \psi, a, d) = \{x \in P : a \leq \psi(x), \gamma(x) \leq d\}$.

Next, we list the fixed point theorem due to Avery-Peterson [21].

Lemma 2.3 (see [21]). *Let P be a cone in a real Banach space E and γ, θ, α, ψ defined as above; moreover, ψ satisfies $\psi(\lambda' x) \leq \lambda' \psi(x)$ for $0 \leq \lambda' \leq 1$ such that for some positive numbers h and d,*

$$\alpha(x) \leq \psi(x),\ \|x\| \leq h\gamma(x), \tag{2.2}$$

for all $x \in \overline{P(\gamma, d)}$. Suppose that $A : \overline{P(\gamma, d)} \to \overline{P(\gamma, d)}$ is completely continuous and there exist positive real numbers a, b, c with $a < b$ such that

(i) $\{x \in P(\gamma, \theta, \alpha, b, c, d) : \alpha(x) > b\} \neq \emptyset$ *and* $\alpha(A(x)) > b$ *for* $x \in P(\gamma,\ \theta,\ \alpha,\ b,\ c,\ d)$,

(ii) $\alpha(A(x)) > b$ *for* $x \in P(\gamma,\ \alpha,\ b,\ d)$ *with* $\theta(A(x)) > c$,

(iii) $0 \notin R(\gamma, \psi, a, d)$ *and* $\psi(A(x)) < a$ *for all* $x \in R(\gamma,\ \psi,\ a,\ d)$ *with* $\psi(x) = a$.

Then, A has at least three fixed points $x_1,\ x_2,\ x_3 \in \overline{P(\gamma, d)}$ such that

$$\gamma(x_i) \leq d \quad \text{for } i = 1, 2, 3, b < \alpha(x_1),\ a < \psi(x_2),\ \alpha(x_2) < b \text{ with } \psi(x_3) < a. \tag{2.3}$$

to Contents, page 1

Now, let $E = C^1([0,1],\mathbb{R})$. Then, E is a Banach space with norm

$$\|u\| = \max\left\{\max_{t\in[0,1]}|u(t)|, \max_{t\in[0,1]}\left|u'(t)\right|\right\}. \tag{2.4}$$

Define a cone $P \subset E$ by

$$P = \left\{u \in E \mid u(0) = 0, u \text{ is concave, nonnegative on } [0,1] \text{and } u \text{ is symmetricon} [\eta, 1]\right\}. \tag{2.5}$$

The following lemma can be founded in [11], which is necessary to prove our result.

Lemma 2.4 (see [11]). *If $u \in P$, then the following statements are true:*

(i) $u(t) \geq (u(\omega_1)/\omega_1)\min\{t, 1+\eta-t\}$ *for* $t \in [0,1]$, *here* $\omega_1 = (\eta+1)/2$,

(ii) $u(t) \geq (\eta/\omega_1)u(\omega_1)$ *for* $t \in [\eta, \omega_1]$,

(iii) $\max_{t\in[0,1]} u(t) = u(\omega_1)$.

Lemma 2.5. *If $u \in P$, then the following statements are true:*

(i) $u(t) \leq \max_{t\in[0,1]}|u'(t)|$,

(ii) $\|u(t)\| = \max_{t\in[0,1]}|u'(t)| = \max\{|u'(0)|, |u'(1)|\}$,

(iii) $\min_{t\in[0,\omega_1]} u(t) = u(0)$ *and* $\min_{t\in[\omega_1,1]} u(t) = u(1)$.

Proof. (i) Since

$$u(t) = u(0) + \int_0^t u'(t)dt \quad \text{for } t \in [0,1], \tag{2.6}$$

which reduces to

$$u(t) \leq \int_0^t \left|u'(t)\right|dt \leq \max_{t\in[0,1]}\left|u'(t)\right|. \tag{2.7}$$

(ii) By using $u''(t) \leq 0$ for $t \in [0,1]$, we have $u'(t)$ is monotone decreasing function on $[0,1]$. Moreover,

$$\max_{t\in[0,1]} u(t) = u\left(\frac{\eta+1}{2}\right) = u(\omega_1), \tag{2.8}$$

which implies that $u'(\omega_1) = 0$, so, $u'(t) \geq 0$ for $t \in [0,\omega_1]$ and $u'(t) \leq 0$ for $t \in [\omega_1, 1]$. □

Now, we define the operator $A : P \to E$ by

$$(Au)(t) = \begin{cases} \displaystyle\int_0^t \varphi_q\left(\int_s^{\omega_1} h(r)f(r,u(r),u'(r))dr\right)ds & \text{for } t \in [0,\omega_1], \\ \displaystyle w(\eta) + \int_t^1 \varphi_q\left(\int_{\omega_1}^s h(r)f(r,u(r),u'(r))dr\right)ds & \text{for } t \in [\omega_1,1], \end{cases} \tag{2.9}$$

here, $w(\eta) = (Au)(\eta)$.

Lemma 2.6. *$A : P \to P$ is a completely continuous operator.*

Proof. In fact, $(Au)(t) \geq 0$ for $t \in [0,1]$, $(Au)(\eta) = (Au)(1)$ and $(Au)(0) = 0$.

It is easy to see that the operator A is pseudosymmetric about ω_1 on $[0,1]$.

In fact, for $t \in [\eta,\omega_1]$, we have $1-t+\eta \in [\omega_1,1]$, and according to the integral transform, one has

$$\begin{aligned} &\int_{1-t+\eta}^1 \varphi_q\left(\int_{\omega_1}^s h(r)f(r,u(r),u'(r))dr\right)ds \\ &\quad= \int_\eta^t \varphi_q\left(\int_{s_1}^{\omega_1} h(r_1)f(r_1,u(r_1),u'(r_1))dr_1\right)ds_1, \end{aligned} \tag{2.10}$$

here, $s = 1 - s_1 + \eta,\ r = 1 - r_1 + \eta$. Hence,

$$\begin{aligned} (Au)(1-t+\eta) &= w(\eta) + \int_{1-t+\eta}^1 \varphi_q\left(\int_{\omega_1}^s h(r)f(r,u(r),u'(r))dr\right)ds \\ &= w(\eta) + \int_\eta^t \varphi_q\left(\int_{s_1}^{\omega_1} h(r_1)f(r_1,u(r_1),u'(r_1))dr_1\right)ds_1 \\ &= \int_0^\eta \varphi_q\left(\int_s^{\omega_1} h(r)f(r,u(r),u'(r))dr\right)ds \\ &\quad+ \int_\eta^t \varphi_q\left(\int_s^{\omega_1} h(r)f(r,u(r),u'(r))dr\right)ds \\ &= \int_0^t \varphi_q\left(\int_s^{\omega_1} h(r)f(r,u(r),u'(r))dr\right)ds = (Au)(t). \end{aligned} \tag{2.11}$$

For $t \in [\omega_1,1]$, we note that $1-t+\eta \in [\eta,\omega_1]$, by using the integral transform, one has

$$\begin{aligned} &\int_\eta^{1-t+\eta} \varphi_q\left(\int_s^{\omega_1} h(r)f(r,u(r),u'(r))dr\right)ds \\ &\quad= \int_t^1 \varphi_q\left(\int_{\omega_1}^{s_1} h(r_1)f(r_1,u(r_1),u'(r_1))dr_1\right)ds_1, \end{aligned} \tag{2.12}$$

to Contents, page 1

where $s = 1 - s_1 + \eta, r = 1 - r_1 + \eta$. Thus,

$$
\begin{aligned}
(Au)(1-t+\eta) &= \int_0^{1-t+\eta} \varphi_q\left(\int_s^{\omega_1} h(r) f(r, u(r), u'(r)) dr\right) ds \\
&= w(\eta) + \int_\eta^{1-t+\eta} \varphi_q\left(\int_s^{\omega_1} h(r) f(r, u(r), u'(r)) dr\right) ds \\
&= w(\eta) + \int_t^1 \varphi_q\left(\int_{\omega_1}^{s_1} h(r_1) f(r_1, u(r_1), u'(r_1)) dr_1\right) ds_1 \\
&= w(\eta) + \int_t^1 \varphi_q\left(\int_{\omega_1}^{s} h(r) f(r, u(r), u'(r)) dr\right) ds = (Au)(t).
\end{aligned}
\tag{2.13}
$$

Hence, A is pseudosymmetric about η on $[0,1]$.
In addition,

$$
(Au)'(t) = \varphi_q\left(\int_t^{\omega_1} h(r) f(r, u(r), u'(r)) ds\right) \geq 0, t \in [0, \omega_1] \tag{2.14}
$$

is continuous and nonincreasing in $[0, \omega_1]$; moreover, $\varphi_q(x)$ is a monotone increasing continuously differentiable function

$$
\left(\int_t^{\omega_1} h(s) f(s, u(s), u'(s)) ds\right)' = -h(t) f(t, u(t), u'(t)) \leq 0, t \in [0, \omega_1], \tag{2.15}
$$

it is easy to obtain $(Au)''(t) \leq 0$ for $t \in [0, \omega_1]$. By using the similar way, we can deduce $(Au)''(t) \leq 0$ for $t \in [\omega_1, 1]$. So, $A : P \to P$. It is easy to obtain that $A : P \to P$ is completely continuous. □

Hence, the solutions of BVPs (1.2) are fixed points of the completely continuous operator A.

3. One Solutions

In this section, we will study the existence of one positive pseudosymmetric solution to problem (1.2) by Krasnosel'skii's fixed point theorem in a cone.

to Contents, page 1

Motivated by the notations in reference [23], for $u \in P$, let

$$\begin{aligned} f^0 &= \sup_{t\in[0,1]} \lim_{(u,u')\to(0,0)} \frac{f(t,u,u')}{\varphi_p(|u'|)}, \\ f_0 &= \inf_{t\in[0,1]} \lim_{(u,u')\to(0,0)} \frac{f(t,u,u')}{\varphi_p(|u'|)}, \\ f^\infty &= \sup_{t\in[0,1]} \lim_{(u,u')\to(\infty,\infty)} \frac{f(t,u,u')}{\varphi_p(|u'|)}, \\ f_\infty &= \inf_{t\in[0,1]} \lim_{(u,u')\to(\infty,\infty)} \frac{f(t,u,u')}{\varphi_p(|u'|)}. \end{aligned} \tag{3.1}$$

In the following, we discuss the problem (1.2) under the following four possible cases.

Theorem 3.1. *If $f^0 = 0$ and $f_\infty = \infty$, problem* (1.2) *has at least one positive pseudosymmetric solution u.*

Proof. In view of $f^0 = 0$, there exists an $H_1 > 0$ such that

$$f(t,u,u') \le \varphi_p(\varepsilon)\varphi_p(|u'|) = \varphi_p(\varepsilon|u'|) \quad \text{for } (t,u,u') \in [0,1]\times(0,H_1]\times[-H_1,H_1], \tag{3.2}$$

here, $\varepsilon > 0$ and satisfies

$$\varepsilon\varphi_q\left(\int_0^{\omega_1} h(s)ds\right) \le 1. \tag{3.3}$$

If $u \in P$ with $\|u\| = H_1$, by Lemma 2.5, we have

$$u(t) \le \max_{t\in[0,1]}|u'(t)| \le \|u\| = H_1 \quad \text{for } t \in [0,1], \tag{3.4}$$

hence,

$$\begin{aligned} \|Au\| &= \max\{|(Au)'(0)|, |(Au)'(1)|\} \\ &= \max\left\{\varphi_q\left(\int_0^{\omega_1} h(r)f(r,u(r),u'(r))dr\right), \varphi_q\left(\int_{\omega_1}^1 h(r)f(r,u(r),u'(r))dr\right)\right\} \\ &\le \varepsilon\max_{t\in[0,1]}|u'(t)|\varphi_q\left(\int_0^{\omega_1} h(s)ds\right) \le \|u\|. \end{aligned} \tag{3.5}$$

If set $\Omega_{H_1} = \{u \in E : \|u\| < H_1\}$, one has $\|Au\| \le \|u\|$ for $u \in P \cap \partial\Omega_{H_1}$.

According to $f_\infty = \infty$, there exists an $H_2' > 0$ such that

$$f(t,u,u') \ge \max_{t\in[0,1]}\varphi_p(k)\varphi_p(|u'|) = \max_{t\in[0,1]}\varphi_p(k|u'|), \tag{3.6}$$

to Contents, page 1

where $(t,u,u') \in [0,1] \times [H_2', \infty) \times (-\infty, H_2'] \cup [H_2', \infty)$, $k > 0$ and satisfies

$$k\varphi_q\left(\int_{\omega_1}^{1} h(r)dr\right) \geq 1. \tag{3.7}$$

Set

$$\begin{gathered} H_2 = \max\left\{2H_1, \frac{\omega_1}{\eta}H_2'\right\}, \ \Omega_{H_2^*} = \{u \in E : \|u\| < 5H_2\}, \\ \Omega_{H_2} = \left\{u \in \Omega_{H_2^*} : u(\omega_1) < H_2\right\}. \end{gathered} \tag{3.8}$$

For $u \in P \cap \partial\Omega_{H_2}$, we have $u(\omega_1) = H_2$ since $u(t) \leq |u'(t)|$ for $u \in P$.
If $u \in P$ with $u(\omega_1) = H_2$, Lemmas 2.4 and 2.5 reduce to

$$\min_{t\in[\omega_1,1]}\left|u'(t)\right| \geq \min_{t\in[\omega_1,1]} u(t) = u(1) \geq \frac{\eta u(\omega_1)}{\omega_1} \geq H_2'. \tag{3.9}$$

For $u \in P \cap \partial\Omega_{H_2}$, according to (3.6), (3.7) and (3.9), we get

$$\begin{aligned} \|Au\| &= \max\left\{\varphi_q\left(\int_0^{\omega_1} h(r)f(r,u(r),u'(r))dr\right), \varphi_q\left(\int_{\omega_1}^{1} h(r)f(r,u(r),u'(r))dr\right)\right\} \\ &\geq \varphi_q\left(\int_{\omega_1}^{1} h(r)f(r,u(r),u'(r))dr\right) \\ &\geq k\max_{t\in[0,1]}\left|u'(t)\right|\varphi_q\left(\int_1^{\omega_1} h(r)dr\right) = \|u\|. \end{aligned} \tag{3.10}$$

Thus, by (i) of Lemma 2.2, the problem (1.2) has at least one positive pseudosymmetric solution u in $P \cap (\overline{\Omega}_{H_2} \setminus \Omega_{H_1})$. □

Theorem 3.2. *If $f_0 = \infty$ and $f^\infty = 0$, problem* (1.2) *has at least one positive pseudosymmetric solution u.*

Proof. Since $f_0 = \infty$, there exists an $H_3 > 0$ such that

$$f(t,u,u') \geq \max_{t\in[0,1]}\varphi_p(m)\varphi_p\left(\left|u'\right|\right) = \max_{t\in[0,1]}\varphi_p\left(m\left|u'\right|\right), \tag{3.11}$$

here, $(t,u,u') \in [0,1] \times (0,H_3] \times [-H_3,H_3]$ and m is such that

$$m\varphi_q\left(\int_{\omega_1}^{1} h(r)dr\right) \geq 1. \tag{3.12}$$

to Contents, page 1

If $u \in P$ with $\|u\| = H_3$, Lemma 2.5 implies that

$$u(t) \le \max_{t\in[0,1]} \left|u'(t)\right| \le \|u\| = H_3 \quad \text{for } t \in [0,1], \tag{3.13}$$

now, by (3.11), (3.12), and (3.13), we have

$$\begin{aligned} \|Au\| &= \max\left\{\varphi_q\left(\int_0^{\omega_1} h(r)f(r,u(r),u'(r))dr\right), \varphi_q\left(\int_{\omega_1}^1 h(r)f(r,u(r),u'(r))dr\right)\right\} \\ &\ge \varphi_q\left(\int_{\omega_1}^1 h(r)f(r,u(r),u'(r))dr\right) \ge m\max_{t\in[0,1]}\left|u'(t)\right|\varphi_q\left(\int_{\omega_1}^1 h(r)dr\right) = \|u\|. \end{aligned} \tag{3.14}$$

If let $\Omega_{H_3} = \{u \in E : \|u\| < H_3\}$, one has $\|Au\| \ge \|u\|$ for $u \in P \cap \partial\Omega_{H_3}$.

Now, we consider $f^\infty = 0$.

Suppose that f is bounded, for some constant $K > 0$, then

$$f(t,u,u') \le \varphi_p(K) \quad \forall (t,u,u') \in [0,1] \times [0,\infty) \times (-\infty,\infty). \tag{3.15}$$

Pick

$$H_4 \ge \max\left\{H_4', 2H_3,\ K\varphi_q\left(\int_0^{\omega_1} h(s)ds\right), \frac{C}{\delta}\right\}, \tag{3.16}$$

here, C is an arbitrary positive constant and satisfy the (3.21). Let

$$\Omega_{H_4} = \{u \in E : \|u\| < H_4\}. \tag{3.17}$$

If $u \in P \cap \partial\Omega_{H_4}$, one has $\|u\| = H_4$, then (3.15) and (3.16) imply that

$$\begin{aligned} \|Au\| &= \max\left\{\varphi_q\left(\int_0^{\omega_1} h(r)f(r,u(r),u'(r))dr\right), \varphi_q\left(\int_{\omega_1}^1 h(r)f(r,u(r),u'(r))dr\right)\right\} \\ &\le K\varphi_q\left(\int_0^{\omega_1} h(s)ds\right) \le H_4 = \|u\|. \end{aligned} \tag{3.18}$$

Suppose that f is unbounded.

By definition of $f^\infty = 0$, there exists $H_4' > 0$ such that

$$f(t,u,u') \le \varphi_p(\delta)\varphi_p(\left|u'\right|) = \varphi_p(\delta\left|u'\right|), \tag{3.19}$$

where $(t,u,u') \in [0,\omega_1] \times [H_4',\infty) \times (-\infty, H_4'] \cup [H_4',\infty)$ and $\delta > 0$ satisfies

$$\delta\varphi_q\left(\int_0^{\omega_1} h(s)ds\right) \le 1. \tag{3.20}$$

to Contents, page 1

From $f \in C([0,1] \times [0,+\infty) \times (-\infty,\infty), [0,+\infty))$, we have

$$f(t,u,u') \le \varphi_p(C) \quad \text{for } (t,u,u') \in [0,1] \times [0,H_4'] \times [-H_4', H_4'], \tag{3.21}$$

here, C is an arbitrary positive constant.

Then, for $(t,u,u') \in [0,1] \times [0,\infty) \times (-\infty,\infty)$, we have

$$f(t,u,u') \le \max\{\varphi_p(C), \varphi_p(\delta)\varphi_p(|u'|)\}. \tag{3.22}$$

If $u \in P \cap \partial\Omega_{H_4}$, one has $\|u\| = H_4$, which reduces to

$$\begin{aligned} \|Au\| &= \max\left\{\varphi_q\left(\int_0^{\omega_1} h(r)f(r,u(r),u'(r))dr\right), \varphi_q\left(\int_{\omega_1}^1 h(r)f(r,u(r),u'(r))dr\right)\right\} \\ &\le \max\{C, \delta\|u'\|\}\varphi_q\left(\int_0^{\omega_1} h(r)dr\right) \\ &\le H_4 = \|u\|. \end{aligned} \tag{3.23}$$

Consequently, for any cases, if we take $\Omega_{H_4} = \{u \in E : \|u\| < H_4\}$, we have $\|Au\| \le \|u\|$ for $u \in P \cap \partial\Omega_{H_4}$. Thus, the condition (ii) of Lemma 2.2 is satisfied.

Consequently, the problem (1.2) has at least one positive pseudosymmetric solution

$$u \in P \cap \left(\overline{\Omega}_{H_4} \setminus \Omega_{H_3}\right) \text{ with } H_3 \le \|u\| \le H_4. \tag{3.24}$$

□

Theorem 3.3. *Suppose that the following conditions hold:*

(i) *there exist nonzero finite constants c_1 and c_2 such that $f^0 = c_1$ and $f_\infty = c_2$,*

(ii) *there exist nonzero finite constants c_3 and c_4 such that $f_0 = c_3$ and $f^\infty = c_4$.*

Then, problem (1.2) *has at least one positive pseudosymmetric solution u.*

Proof. (i) In view of $f^0 = c_1$, there exists an $H_5 > 0$ such that

$$\begin{aligned} f(t,u,u') &\le \varphi_p(\varepsilon + c_{11})\varphi_p(|u'|) \\ &= \varphi_p((\varepsilon + c_{11})|u'|) \quad \text{for } (t,u,u') \in [0,1] \times (0,H_5] \times [-H_5, H_5], \end{aligned} \tag{3.25}$$

here, $c_1 = \varphi_p(c_{11} + \varepsilon)$, $\varepsilon > 0$ and satisfies

$$(\varepsilon + c_{11})\varphi_q\left(\int_0^{\omega_1} h(s)ds\right) \le 1. \tag{3.26}$$

If $u \in P$ with $\|u\| = H_5$, by Lemma 2.5, we have

$$u(t) \le |u'(t)| \le \|u\| = H_5 \quad \text{for } t \in [0,1], \tag{3.27}$$

to Contents, page 1

hence,

$$\begin{aligned}\|Au\| &= \max\{|(Au)'(0)|, |(Au)'(1)|\}\\ &= \max\left\{\varphi_q\left(\int_0^{\omega_1} h(r)f(r,u(r),u'(r))dr\right), \varphi_q\left(\int_{\omega_1}^1 h(r)f(r,u(r),u'(r))dr\right)\right\}\\ &\le (\varepsilon + c_{11})\max_{t\in[0,1]}|u'(t)|\varphi_q\left(\int_0^{\omega_1} h(s)ds\right) \le \|u\|.\end{aligned} \tag{3.28}$$

If set $\Omega_{H_5} = \{u \in E : \|u\| < H_5\}$, one has $\|Au\| \le \|u\|$ for $u \in P \cap \partial\Omega_{H_5}$.

According to $f_\infty = c_2$, there exists an $H_6' > 0$ such that

$$f(t,u,u') \ge \max_{t\in[0,1]}\varphi_p(c_{22}-\varepsilon)\varphi_p(|u'|) = \max_{t\in[0,1]}\varphi_p((c_{22}-\varepsilon)|u'|), \tag{3.29}$$

where $(t,u,u') \in [0,1] \times [H_6',\infty) \times (-\infty, H_6'] \cup [H_6',\infty)$, $c_2 = \varphi_p(c_{22}-\varepsilon)$, $\varepsilon > 0$ and satisfies

$$(c_{22}-\varepsilon)\varphi_q\left(\int_{\omega_1}^1 h(r)dr\right) \ge 1. \tag{3.30}$$

Set

$$\begin{gathered} H_6 = \max\left\{2H_5, \frac{\omega_1}{\eta}H_6'\right\}, \qquad \Omega_{H_6^*} = \{u \in E : \|u\| < 5H_6\},\\ \Omega_{H_6} = \left\{u \in \Omega_{H_6^*} : u(\omega_1) < H_6\right\}.\end{gathered} \tag{3.31}$$

If $u \in P$ with $u(\omega_1) = H_6$, Lemmas 2.4 and 2.5 reduce to

$$\min_{t\in[\omega_1,1]}|u'(t)| \ge \min_{t\in[\omega_1,1]} u(t) = u(1) \ge \frac{\eta u(\omega_1)}{\omega_1} \ge H_6'. \tag{3.32}$$

For $u \in P \cap \partial\Omega_{H_6}$, according to (3.29), (3.30) and (3.32), we get

$$\begin{aligned}\|Au\| &= \max\left\{\varphi_q\left(\int_0^{\omega_1} h(r)f(r,u(r),u'(r))dr\right), \varphi_q\left(\int_{\omega_1}^1 h(r)f(r,u(r),u'(r))dr\right)\right\}\\ &\ge \varphi_q\left(\int_{\omega_1}^1 h(r)f(r,u(r),u'(r))dr\right)\\ &\ge (c_{22}-\varepsilon)\max_{t\in[0,1]}|u'(t)|\varphi_q\left(\int_1^{\omega_1} h(r)dr\right) = \|u\|.\end{aligned} \tag{3.33}$$

Thus, by (i) of Lemma 2.2, the problem (1.2) has at least one positive pseudosymmetric solution u in $P \cap (\overline{\Omega}_{H_6} \setminus \Omega_{H_5})$.

(ii) By using the similar way as to Theorem 3.2, we can prove to it. $\square$

to Contents, page 1

4. Triple Solutions

In the previous section, some results on the existence of at least one positive pseudosymmetric solutions to problem (1.2) are obtained. In this section, we will further discuss the existence criteria for at least *three* and arbitrary odd positive pseudosymmetric solutions of problems (1.2) by using the Avery-Peterson fixed point theorem [21].

Choose a $r \in (\eta, \omega_1)$, for the notational convenience, we denote

$$M = \omega_1 \varphi_q\left(\int_0^{\omega_1} h(r)dr\right), \qquad N = \eta\varphi_q\left(\int_\eta^{\omega_1} h(r)dr\right), \qquad W = \varphi_q\left(\int_0^{\omega_1} h(r)dr\right). \tag{4.1}$$

Define the nonnegative continuous convex functionals θ and γ, nonnegative continuous concave functional α, and nonnegative continuous functional φ, respectively, on P by

$$\begin{aligned} \gamma(u) &= \max_{t\in[0,1]}\left|u'(t)\right| = \max\{u'(0), u'(1)\} = \|u\|, \\ \psi(u) = \theta(u) &= \max_{t\in[0,\omega_1]} u(t) = u(\omega_1) \le \|u\|, \\ \alpha(u) &= \min_{t\in[\eta,\omega_1]} u(t) = u(\eta). \end{aligned} \tag{4.2}$$

Now, we state and prove the results in this section.

Theorem 4.1. *Suppose that there exist constants a^*, b^*, and d^* such that $0 < a^* < b^* < (N/W)d^*$. In addition, f satisfies the following conditions:*

(i) $f(t,u,u') \le \varphi_p(d^*/W)$ *for* $(t,u,u') \in [0,1]\times[0,d^*]\times[-d^*,d^*]$,

(ii) $f(t,u,u') > \varphi_p(b^*/N)$ *for* $(t,u,u') \in [\eta,\omega_1]\times[b^*,d^*]\times[-d^*,d^*]$,

(iii) $f(t,u,u') < \varphi_p(a^*/M)$ *for* $(t,u,u') \in [0,\omega_1]\times[0,a^*]\times[-d^*,d^*]$.

Then, problem (1.2) *has at least three positive pseudosymmetric solutions u_1, u_2, and u_3 such that*

$$\begin{aligned} &\|x_i\| \le d^* \quad \text{for } i = 1,2,3,\ b^* < \min_{t\in[\eta,\omega_1]} u_1(t),\ a^* < \max_{t\in[0,1]} u_2(t), \\ &\min_{t\in[\eta,\omega_1]} u_2(t) < b^* \quad \text{with} \max_{t\in[0,1]} u_3(t) < a^*. \end{aligned} \tag{4.3}$$

Proof. According to the definition of completely continuous operator A and its properties, we need to show that all the conditions of Lemma 2.3 hold with respect to A.

It is obvious that

$$\begin{aligned} \psi(\lambda' u) &= \lambda' u(\omega_1) = \lambda'\psi(u) \quad \text{for } 0 < \lambda' < 1, \\ \alpha(u) &\le \psi(u) \quad \forall u \in P, \\ \|u\| &= \gamma(u) \quad \forall u \in P. \end{aligned} \tag{4.4}$$

to Contents, page 1

Firstly, we show that $A : \overline{P(\gamma, d^*)} \to \overline{P(\gamma, d^*)}$.

For any $u \in \overline{P(\gamma, d^*)}$, we have

$$u(t) \le \max_{t\in[0,1]} |u'(t)| \le \|u\| = \gamma(u) \le d^* \quad \text{for } t \in [0,1], \tag{4.5}$$

hence, the assumption (i) implies that

$$\begin{aligned} \|Au\| &= \max\left\{\varphi_q\left(\int_0^{\omega_1} h(r) f(r, u(r), u'(r)) dr\right), \varphi_q\left(\int_{\omega_1}^1 h(r) f(r, u(r), u'(r)) dr\right)\right\} \\ &\le \frac{d^*}{W}\varphi_q\left(\int_0^{\omega_1} h(r) dr\right) = d^*. \end{aligned} \tag{4.6}$$

From the above analysis, it remains to show that (i)–(iii) of Lemma 2.3 hold.

Secondly, we verify that condition (i) of Lemma 2.3 holds; let $u(t) \equiv (tb^*/\eta) + b^*, t \in [0,1]$, and it is easy to see that

$$\begin{aligned} \alpha(u) &= u(\eta) = 2b^* > b^*, \\ \theta(u) &= u(\omega_1) = \frac{\omega_1 b^*}{\eta} + b^* \le \frac{\omega_1 b^*}{\eta} + b^*, \end{aligned} \tag{4.7}$$

in addition, we have $\gamma(u) = (b^*/\eta) < d^*$, since $b^* < (N/W)d^*$. Thus

$$\left\{u \in P\left(\gamma, \theta, \alpha, b^*, \frac{\omega_1 b^*}{\eta} + b^*, d^*\right) : \alpha(x) > b^*\right\} \ne \emptyset. \tag{4.8}$$

For any

$$u \in P\left(\gamma, \theta, \alpha, b^*, \frac{\omega_1 b^*}{\eta} + b^*, d^*\right), \tag{4.9}$$

one has

$$b^* \le u(t) \le \|u\| \le d^* \quad \forall t \in [\eta, \omega_1], \tag{4.10}$$

it follows from the assumption (ii) that

$$\begin{aligned} \alpha(Au) = (Au)(\eta) &= \int_0^\eta \varphi_q\left(\int_s^{\omega_1} h(r) f(r, u(r), u'(r)) dr\right) ds \\ &\ge \int_0^\eta \varphi_q\left(\int_\eta^{\omega_1} h(r) f(r, u(r), u'(r)) dr\right) ds \\ &> \frac{b^*}{N}\eta\varphi_q\left(\int_\eta^{\omega_1} h(r) dr\right) = b^*. \end{aligned} \tag{4.11}$$

to Contents, page 1

Thirdly, we prove that the condition (ii) of Lemma 2.3 holds. In fact,

$$\begin{gathered}\alpha(Au) = Au(\eta), \\ \theta(Au) = \max_{t\in[0,\omega_1]} A(u) = Au(\omega_1).\end{gathered} \tag{4.12}$$

For any $u \in P(\gamma,\alpha,b^*,d^*)$ with $\theta(Au) > (\omega_1 b^*/\eta) + b^*$, we have

$$\alpha(Au) = Au(\eta) \geq \frac{\eta}{\omega_1} Au(\omega_1) \geq \frac{\eta}{\omega_1}\theta(Au) = b^* + \frac{\omega_1 b^*}{\eta} > b^*. \tag{4.13}$$

Finally, we check condition (iii) of Lemma 2.3.

Clearly, since $\psi(0) = 0 < a^*$, we have $0 \notin R(\gamma,\psi,a^*,d^*)$. If

$$u \in R(\gamma,\psi,a^*,d^*) \text{ with } \psi\ (u) = \max_{t\in[0,\omega_1]} u(t) = u(\omega_1) = a^*, \tag{4.14}$$

then

$$\begin{gathered}0 \leq u(t) \leq a^* \quad \forall t \in [0,\omega_1], \\ \max_{t\in[0,1]}\left|u'(t)\right| = \|u\| = \gamma(u) \leq d^*.\end{gathered} \tag{4.15}$$

Hence, by assumption (iii), we have

$$\begin{aligned}\psi(Au) &= (Au)(\omega_1) \\ &\leq \int_0^{\omega_1} \varphi_q\left(\int_0^{\omega_1} h(r) f(r,u(r),u'(r))dr\right)ds \\ &< \frac{a^*}{M}\omega_1\varphi_q\left(\int_0^{\omega_1} h(r)dr\right) = a^*.\end{aligned} \tag{4.16}$$

Consequently, from above, all the conditions of Lemma 2.3 are satisfied. The proof is completed. □

Corollary 4.2. *If the condition* (i) *in Theorem 4.1 is replaced by the following condition (i′):*

(i′) $\lim_{(u,u')\to(\infty,\infty)} (f(t,u,u')/(\varphi_p(|u'|))) \leq \varphi_p(1/W)$,

then the conclusion of Theorem 4.1 also holds.

Proof. From Theorem 4.1, we only need to prove that (i′) implies that (i) holds. That is, assume that (i′) holds, then there exists a number $d^* \geq (W/N)b^*$ such that

$$f(t,u,u') \leq \varphi_p\left(\frac{d^*}{W}\right) \quad \text{for } (t,u,u') \in [0,1]\times[0,d^*]\times[-d^*,d^*]. \tag{4.17}$$

to Contents, page 1

Suppose on the contrary that for any $d^* \geq (W/N)b^*$, there exists $(u_c, u'_c) \in [0, d^*] \times [-d^*, d^*]$ such that

$$f(t, u_c, u'_c) > \varphi_p\left(\frac{d^*}{W}\right) \quad \text{for } t \in [0,1]. \tag{4.18}$$

Hence, if we choose $c_n^* > (W/N)b^*$ $(n = 1, 2, \ldots)$ with $c_n^* \to \infty$, then there exist $(u_n, u'_n) \in [0, c_n^*] \times [-c_n^*, c_n^*]$ such that

$$f(t, u_n, u'_n) > \varphi_p\left(\frac{c_n^*}{W}\right) \quad \text{for } t \in [0,1], \tag{4.19}$$

and so

$$\lim_{n\to\infty} f(t, u_n, u'_n) = \infty \quad \text{for } t \in [0,1]. \tag{4.20}$$

Since the condition (i′) holds, there exists $\tau > 0$ satisfying

$$f(t, u, u') \leq \varphi_p\left(\frac{|u'|}{W}\right) \quad \text{for } (t, u, u') \in [0,1] \times [\tau, \infty) \times (-\infty, \tau] \cup [\tau, \infty). \tag{4.21}$$

Hence, we have

$$u_n < |u'_n| \leq \tau. \tag{4.22}$$

Otherwise, if

$$|u'_n| > u_n > \tau \quad \text{for } t \in [0,1], \tag{4.23}$$

it follows from (4.21) that

$$f(t, u_n, u'_n) \leq \varphi_p\left(\frac{u_n}{W}\right) \leq \varphi_p\left(\frac{c_n^*}{W}\right) \quad \text{for } t \in [0,1], \tag{4.24}$$

which contradicts (4.19).

Let

$$W = \max_{(t,u,u') \in [0,1]\times[0,\tau]\times[-\tau,\tau]} f(t, u, u'), \tag{4.25}$$

then

$$f(t, u_n, u'_n) \leq W (n = 1, 2, \ldots), \tag{4.26}$$

which also contradicts (4.20). □

to Contents, page 1

Theorem 4.3. *Suppose that there exist constants a_i^*, b_i^*, and d_i^* such that*

$$0 < a_1^* < b_1^* < \frac{N}{W} d_2^* < a_2^* < b_2^* < \frac{N}{W} d_3^* < \cdots < a_n^* < b_n^* < \frac{N}{W} d_{n+1}^*, \tag{4.27}$$

here, $n \in \mathbb{N}$ and $i = 1, 2, \ldots, n$. In addition, suppose that f satisfies the following conditions:

(i) $f(t, u, u') \le \varphi_p(d_i^*/W)$ *for* $(t, u, u') \in [0,1] \times [0, d_i^*] \times [-d_i^*, d_i^*]$,

(ii) $f(t, u, u') > \varphi_p(b_i^*/N)$ *for* $(t, u, u') \in [\eta, \omega_1] \times [b_i^*, d_i^*] \times [-d_i^*, d_i^*]$,

(iii) $f(t, u, u') < \varphi_p(a_i^*/M)$ *for* $(t, u, u') \in [0, \omega_1] \times [0, a_i^*] \times [-d_i^*, d_i^*]$.

Then, problem (1.2) *has at least* $2n - 1$ *positive pseudosymmetric solutions.*

Proof. When $n = 1$, it is immediate from condition (i) that

$$A : \overline{P}_{a_1^*} \longrightarrow P_{a_1^*} \subset \overline{P}_{a_1^*}. \tag{4.28}$$

It follows from the Schauder fixed point theorem that A has at least one fixed point

$$u_1 \in \overline{P}_{a_1^*}, \tag{4.29}$$

which means that

$$\|u_1\| \le a_1^*. \tag{4.30}$$

When $n = 2$, it is clear that Theorem 4.1 holds (with $a^* = a_1^*$, $b^* = b_1^*$, and $d^* = d_2^*$). Then, there exists at least three positive pseudosymmetric solutions u_1, u_2, and u_3 such that

$$\begin{gathered} \|x_1\| \le d_2^*, \quad \|x_2\| \le d_2^*, \quad \|x_3\| \le d_2^*, \quad b^* < \min_{t \in [\eta, \omega_1]} u_1(t), \quad a_1^* < \max_{t \in [0,1]} u_2(t), \\ \min_{t \in [\eta, \omega_1]} u_2(t) < b_1^* \text{ with } \max_{t \in [0,1]} u_3(t) < a_1^*. \end{gathered} \tag{4.31}$$

Following this way, we finish the proof by induction. The proof is complete. □

5. Examples

In this section, we present two simple examples to illustrate our results.

Example 5.1. Consider the following BVPs:

$$\begin{gathered} (\varphi_p(u'(t)))' + t\left(t + 1 + |u'(t)|^{p-2}\right) = 0, \quad t \in [0,1], \\ u(0) = 0, \; u(0.2) = u(1). \end{gathered} \tag{5.1}$$

to Contents, page 1

Note that

$$f_0 = \inf_{t\in[0,1]} \lim_{(u,u')\to(0,0)} \frac{t+1+|u'(t)|^{p-2}}{|u'(t)|^{p-2}u'(t)} = \infty,$$
$$f^\infty = \sup_{t\in[0,1]} \lim_{(u,u')\to(\infty,\infty)} \frac{t+1+|u'(t)|^{p-2}}{|u'(t)|^{p-2}u'(t)} = 0. \tag{5.2}$$

Hence, Theorem 3.2 implies that the BVPs in (5.1) have at least one pseudosymmetric solution u.

Example 5.2. Consider the following BVPs with $p = 3$:

$$\left(\varphi_p(u'(t))\right)' + h(t)f(t,u(t),u'(t)) = 0, t \in [0,1],$$
$$u(0) = 0,\ u\ (0.2) = u(1), \tag{5.3}$$

where $h(t) = 2t$ and

$$f(t,u,u') = \begin{cases} t+4+\left(\dfrac{u'}{5.5}\right)^2, & u \in [0,0.9], \\ t+750u-671+\left(\dfrac{u'}{5.5}\right)^2, & u \in [0.9,1], \\ t+79+\left(\dfrac{u'}{5.5}\right)^2, & u \in [1,5.5], \\ t+14.364u+\left(\dfrac{u'}{5.5}\right)^2, & u \in [5.5,+\infty). \end{cases} \tag{5.4}$$

Note that $\eta = 0.2,\ \omega_1 = 0.6$, then a direct calculation shows that

$$M = \omega_1\varphi_q\left(\int_0^{\omega_1} h(r)dr\right) = 0.6 \times 0.6 = 0.36,\ N \approx 0.1131,\ W = 0.6. \tag{5.5}$$

If we take $a' = 0.9,\ b' = 1,\ d' = 5.5$, then $a' < b' < (N/W)d'$ holds; furthermore,

$$f(t,u,u') < 82 < 84.028 \approx \varphi_p\left(\frac{d'}{W}\right) \quad \text{for } (t,u,u') \in [0,0.6]\times[0,5.5]\times[-5.5,5.5],$$
$$f(t,u,u') > 79 > 78.176 \approx \varphi_p\left(\frac{b'}{N}\right) \quad \text{for } (t,u,u') \in [0.6,1]\times[1,5.5]\times[-5.5,5.5], \tag{5.6}$$
$$f(t,u,u') < 6.2 < 6.25 = \varphi_p\left(\frac{a'}{M}\right) \quad \text{for } (t,u,u') \in [0,0.6]\times[0,0.9]\times[-5.5,5.5].$$

to Contents, page 1

By Theorem 4.1, we see that the BVPs in (5.3) have at least *three* positive pseudosymmetric solutions u_1,u_2 and u_3 such that

$$\begin{gathered}\|x_i\| \le 5.5 \quad \text{for } i = 1,\ 2,\ 3,\ 1 < \min_{t\in[0.2,0.6]} u_1(t), 0.9 < \max_{t\in[0,1]} u_2(t), \\ \min_{t\in[0.2,0.6]} u_2(t) < 1 \quad \text{with } \max_{t\in[0,1]} u_3(t) < 0.9.\end{gathered} \tag{5.7}$$

Acknowledgments

Y. H. Su Sponsored by Qing Lan Project and XZIT (no. XKY2010101). X. Yan Supported by Fundamental Research Funds for the Central Universities Grant (no. 2010QNA40).

References

[1] C. Bai, "Triple positive solutions of three-point boundary value problems for fourth-order differential equations," *Computers & Mathematics with Applications*, vol. 56, no. 5, pp. 1364–1371, 2008.
[2] R. Ma and H. Ma, "Positive solutions for nonlinear discrete periodic boundary value problems," *Computers & Mathematics with Applications*, vol. 59, no. 1, pp. 136–141, 2010.
[3] R. Ma and B. Zhu, "Existence of positive solutions for a semipositone boundary value problem on the half-line," *Computers & Mathematics with Applications*, vol. 58, no. 8, pp. 1672–1686, 2009.
[4] J.-P. Sun and Y.-H. Zhao, "Multiplicity of positive solutions of a class of nonlinear fractional differential equations," *Computers & Mathematics with Applications*, vol. 49, no. 1, pp. 73–80, 2005.
[5] D.-B. Wang and W. Guan, "Three positive solutions of boundary value problems for p-Laplacian difference equations," *Computers & Mathematics with Applications*, vol. 55, no. 9, pp. 1943–1949, 2008.
[6] Y. Zhu and J. Zhu, "Existence of multiple positive solutions for nth-order p-Laplacian m-point singular boundary value problems," *Journal of Applied Mathematics and Computing*, vol. 34, no. 1-2, pp. 393–405, 2010.
[7] R. Avery and J. Henderson, "Existence of three positive pseudo-symmetric solutions for a one dimensional discrete p-Laplacian," *Journal of Difference Equations and Applications*, vol. 10, no. 6, pp. 529–539, 2004.
[8] R. I. Avery and J. Henderson, "Existence of three positive pseudo-symmetric solutions for a one-dimensional p-Laplacian," *Journal of Mathematical Analysis and Applications*, vol. 277, no. 2, pp. 395–404, 2003.
[9] M. Feng, X. Zhang, and W. Ge, "Exact number of pseudo-symmetric positive solutions for a p-Laplacian three-point boundary value problems and their applications," *Applied Mathematics and Computing*, vol. 33, no. 1-2, pp. 437–448, 2010.
[10] D. Ji, Y. Yang, and W. Ge, "Triple positive pseudo-symmetric solutions to a four-point boundary value problem with p-Laplacian," *Applied Mathematics Letters*, vol. 21, no. 3, pp. 268–274, 2008.
[11] D.-X. Ma and W.-G. Ge, "Existence and iteration of positive pseudo-symmetric solutions for a three-point second-order p-Laplacian BVP," *Applied Mathematics Letters*, vol. 20, no. 12, pp. 1244–1249, 2007.
[12] J. T. Cho and J.-I. Inoguchi, "Pseudo-symmetric contact 3-manifolds. II. When is the tangent sphere bundle over a surface pseudo-symmetric?" *Note di Matematica*, vol. 27, no. 1, pp. 119–129, 2007.
[13] S. W. Ng and A. D. Rae, "The pseudo symmetric structure of bis(dicyclohexylammonium) bis(oxalatotriphenylstannate)," *Zeitschrift für Kristallographie*, vol. 215, no. 3, pp. 199–204, 2000.
[14] T. Jankowski, "Existence of positive solutions to second order four-point impulsive differential problems with deviating arguments," *Computers & Mathematics with Applications*, vol. 58, no. 4, pp. 805–817, 2009.
[15] X.-F. Li and P.-H. Zhao, "The existence of triple positive solutions of nonlinear four-point boundary value problem with p–Laplacian," *Turkish Journal of Mathematics*, vol. 33, no. 2, pp. 131–142, 2009.
[16] B. Sun and W. Ge, "Successive iteration and positive pseudo-symmetric solutions for a three-point second-order p-Laplacian boundary value problems," *Applied Mathematics and Computation*, vol. 188, no. 2, pp. 1772–1779, 2007.

to Contents, page 1

[17] Y. Wang and W. Gao, "Existence of triple positive solutions for multi-point boundary value problems with a one dimensional p-Laplacian," *Computers & Mathematics with Applications*, vol. 54, no. 6, pp. 793–807, 2007.

[18] F. Xu, L. Liu, and Y. Wu, "Multiple positive solutions of four-point nonlinear boundary value problems for a higher-order p-Laplacian operator with all derivatives," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 9, pp. 4309–4319, 2009.

[19] R. I. Avery, "A generalization of the Leggett-Williams fixed point theorem," *Mathematical Sciences Research Hot-Line*, vol. 3, no. 7, pp. 9–14, 1999.

[20] M. A. Krasnosel'skii, *Positive Solutions of Operator Equations*, P. Noordhoff, Groningen, The Netherlands, 1964.

[21] R. I. Avery and A. C. Peterson, "Three positive fixed points of nonlinear operators on ordered banach spaces," *Computers & Mathematics with Applications*, vol. 42, no. 3-5, pp. 313–322, 2001.

[22] D. J. Guo and V. Lakshmikantham, *Nonlinear Problems in Abstract Cones*, vol. 5 of *Notes and Reports in Mathematics in Science and Engineering*, Academic Press, San Diego, Calif, USA, 1988.

[23] H. Wang, "Positive periodic solutions of functional differential equations," *Journal of Differential Equations*, vol. 202, no. 2, pp. 354–366, 2004.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 760209, 14 pages
doi:10.1155/2011/760209

Research Article

Global Nonexistence of Positive Initial-Energy Solutions for Coupled Nonlinear Wave Equations with Damping and Source Terms

Liang Fei[1,2] and Gao Hongjun[1]

[1] *Jiangsu Provincial Key Laboratory for NSLSCS, School of Mathematical Science, Nanjing Normal University, Nanjing 210046, China*
[2] *Department of Mathematics, Anhui Science and Technology University, Fengyang 233100, Anhui, China*

Correspondence should be addressed to Liang Fei, fliangmath@126.com

Received 26 November 2010; Revised 5 June 2011; Accepted 27 June 2011

Academic Editor: Josef Diblík

This work is concerned with a system of nonlinear wave equations with nonlinear damping and source terms acting on both equations. We prove a global nonexistence theorem for certain solutions with positive initial energy.

1. Introduction

In this paper we study the initial-boundary-value problem

$$\begin{gathered}
u_{tt} - \operatorname{div}\left(g\left(|\nabla u|^2\right)\nabla u\right) + |u_t|^{m-1}u_t = f_1(u,v), \quad (x,t) \in \Omega \times (0,T), \\
v_{tt} - \operatorname{div}\left(g\left(|\nabla v|^2\right)\nabla v\right) + |v_t|^{r-1}v_t = f_2(u,v), \quad (x,t) \in \Omega \times (0,T), \\
u(x,t) = v(x,t) = 0, \quad x \in \partial\Omega \times (0,T), \\
u(x,0) = u_0(x), \qquad u_t(x,0) = u_1(x), \quad x \in \Omega, \\
v(x,0) = v_0(x), \qquad v_t(x,0) = v_1(x), \quad x \in \Omega,
\end{gathered} \tag{1.1}$$

where Ω is a bounded domain in $\mathbb{R}^n$ with a smooth boundary $\partial\Omega$, $m, r \geq 1$, and $f_i(\cdot,\cdot) : \mathbb{R}^2 \to \mathbb{R}$ $(i = 1,2)$ are given functions to be specified later. We assume that g is a function which

to Contents, page 1

satisfies

$$g \in C^1, \quad g(s) > 0, \quad g(s) + 2sg'(s) > 0 \tag{1.2}$$

for $s > 0$.

To motivate our work, let us recall some results regarding $g \equiv 1$. The single-wave equation of the form

$$u_{tt} - \Delta u + h(u_t) = f(u), \quad x \in \Omega,\ t > 0 \tag{1.3}$$

in $\Omega \times (0, \infty)$ with initial and boundary conditions has been extensively studied, and many results concerning global existence, blow-up, energy decay have been obtained. In the absence of the source term, that is, ($f = 0$), it is well known that the damping term $h(u_t)$ assures global existence and decay of the solution energy for arbitrary initial data (see [1]). In the absence of the damping term, the source term causes finite time blow-up of solutions with a large initial data (negative initial energy) (see [2, 3]). The interaction between the damping term and the source term makes the problem more interesting. This situation was first considered by Levine [4, 5] in the linear damping case $h(u_t) = au_t$ and a polynomial source term of the form $f(u) = b|u|^{p-2}u$. He showed that solutions with negative initial energy blow up in finite time. The main tool used in [4, 5] is the "concavity method." Georgiev and Todorova in [6] extended Levine's result to the nonlinear damping case $h(u_t) = a|u_t|^{m-2}u_t$. In their work, the authors considered problem (1.3) with $f(u) = b|u|^{p-2}u$ and introduced a method different from the one known as the concavity method and showed that solutions with negative energy continue to exist globally in time if $m \geq p \geq 2$ and blow up in finite time if $p > m \geq 2$ and the initial energy is sufficiently negative. This latter result has been pushed by Messaoudi [7] to the situation where the initial energy $E(0) < 0$ and has been improved by the same author in [8] to accommodate certain solutions with positive initial energy.

In the case of g being a given nonlinear function, the following equation:

$$u_{tt} - g(u_x)_x - u_{xxt} + \delta|u_t|^{m-1}\ u_t = \mu|u|^{p-1}u, \quad x \in (0,1),\ t > 0, \tag{1.4}$$

with initial and boundary conditions has been extensively studied. Equation of type of (1.4) is a class of nonlinear evolution governing the motion of a viscoelastic solid composed of the material of the rate type, see [9–12]. It can also be seen as field equation governing the longitudinal motion of a viscoelastic bar obeying the nonlinear Voigt model, see [13]. In two- and three-dimensional cases, they describe antiplane shear motions of viscoelastic solids. We refer to [14–16] for physical origins and derivation of mathematical models of motions of viscoelastic media and only recall here that, in applications, the unknown u naturally represents the displacement of the body relative to a fixed reference configuration. When $\delta = \mu = 0$, there have been many impressive works on the global existence and other properties of solutions of (1.4), see [9, 10, 17, 18]. Especially, in [19] the authors have proved the global existence and uniqueness of the generalized and classical solution for the initial boundary value problem (1.4) when we replace $\delta|u_t|^{m-1}u_t$ and $\mu|u|^{p-1}u$ by $g(u_t)$ and $f(u)$, respectively. But about the blow-up of the solution for problem, in this paper there has not been any discussion. Chen et al. [20] considered problem (1.4) and first established an

to Contents, page 1

ordinary differential inequality, next given the sufficient conditions of blow-up of the solution of (1.4) by the inequality. In [21], Hao et al. considered the single-wave equation of the form

$$u_{tt} - \operatorname{div}\left(g\left(|\nabla u|^2\right)\nabla u\right) + h(u_t) = f(u), \quad x \in \Omega,\ t > 0 \tag{1.5}$$

with initial and Dirichlet boundary condition, where g satisfies condition (1.2) and

$$g(s) \geq b_1 + b_2 s^q, \quad q \geq 0. \tag{1.6}$$

The damping term has the form

$$h(u_t) = d_1 u_t + d_2 |u_t|^{r-1}\, u_t, \quad r > 1. \tag{1.7}$$

The source term is

$$f(u) = a_1 u + a_2 |u|^{p-1} u \tag{1.8}$$

with $p \geq 1$ for $n = 1, 2$ and $1 \leq n \leq 2n/(n-2)$ for $n \geq 3$, $a_1, a_2, b_1, b_2, d_1, d_2$ are nonnegative constants, and $b_1 + b_2 > 0$. By using the energy compensation method [7, 8, 22], they proved that under some conditions on the initial value and the growth orders of the nonlinear strain term, the damping term, and the source term, the solution to problem (1.5) exists globally and blows up in finite time with negative initial energy, respectively.

Some special cases of system (1.1) arise in quantum field theory which describe the motion of charged mesons in an electromagnetic field, see [23, 24]. Recently, some of the ideas in [6, 22] have been extended to study certain systems of wave equations. Agre and Rammaha [25] studied the system of (1.1) with $g \equiv 1$ and proved several results concerning local and global existence of a weak solution and showed that any weak solution with negative initial energy blows up in finite time, using the same techniques as in [6]. This latter blow-up result has been improved by Said-Houari [26] by considering a larger class of initial data for which the initial energy can take positive values. Recently, Wu et al. [27] considered problem (1.1) with the nonlinear functions $f_1(u, v)$ and $f_2(u, v)$ satisfying appropriate conditions. They proved under some restrictions on the parameters and the initial data several results on global existence of a weak solution. They also showed that any weak solution with initial energy $E(0) < 0$ blows up in finite time.

In this paper, we also consider problem (1.1) and improve the global nonexistence result obtained in [27], for a large class of initial data in which our initial energy can take positive values. The main tool of the proof is a technique introduced by Payne and Sattinger [28] and some estimates used firstly by Vitillaro [29], in order to study a class of a single-wave equation.

to Contents, page 1

2. Preliminaries and Main Result

First, let us introduce some notation used throughout this paper. We denote by $\|\cdot\|_q$ the $L^q(\Omega)$ norm for $1 \le q \le \infty$ and by $\|\nabla \cdot\|_2$ the Dirichlet norm in $H_0^1(\Omega)$ which is equivalent to the $H^1(\Omega)$ norm. Moreover, we set

$$(\varphi, \psi) = \int_\Omega \varphi(x)\psi(x)dx \tag{2.1}$$

as the usual $L^2(\Omega)$ inner product.

Concerning the functions $f_1(u,v)$ and $f_2(u,v)$, we take

$$\begin{aligned} f_1(u,v) &= \left[a|u+v|^{2(p+1)}(u+v) + b|u|^p u|v|^{(p+2)}\right], \\ f_2(u,v) &= \left[a|u+v|^{2(p+1)}(u+v) + b|u|^{(p+2)}|v|^p v\right], \end{aligned} \tag{2.2}$$

where $a, b > 0$ are constants and p satisfies

$$\begin{cases} p > -1, & \text{if } n = 1, 2, \\ -1 < p \le \dfrac{4-n}{n-2}, & \text{if } n \ge 3. \end{cases} \tag{2.3}$$

One can easily verify that

$$uf_1(u,v) + vf_2(u,v) = 2(p+2)F(u,v), \quad \forall (u,v) \in \mathbb{R}^2, \tag{2.4}$$

where

$$F(u,v) = \frac{1}{2(p+2)}\left[a|u+v|^{2(p+2)} + 2b|uv|^{p+2}\right]. \tag{2.5}$$

We have the following result.

Lemma 2.1 (see [30, Lemma 2.1]). *There exist two positive constants c_0 and c_1 such that*

$$\frac{c_0}{2(p+2)}\left(|u|^{2(p+2)} + |v|^{2(p+2)}\right) \le F(u,v) \le \frac{c_1}{2(p+2)}\left(|u|^{2(p+2)} + |v|^{2(p+2)}\right). \tag{2.6}$$

Throughout this paper, we define g by

$$g(s) = b_1 + b_2 s^q, \quad q \ge 0, \; b_1 + b_2 > 0, \tag{2.7}$$

where b_1, b_2 are nonnegative constants. Obviously, g satisfies conditions (1.2) and (1.6). Set

$$G(s) = \int_0^s g(s)ds, \quad s \ge 0. \tag{2.8}$$

to Contents, page 1

In order to state and prove our result, we introduce the following function space:

$$\begin{aligned} Z = \Big\{ (u, v) \mid u, v \in L^\infty\Big([0,T); W_0^{1,2(q+1)}(\Omega) \cap L^{2(p+2)}(\Omega)\Big), \\ u_t \in L^\infty\Big([0,T); L^2(\Omega)\Big) \cap L^{m+1}(\Omega \times (0,T)), \\ v_t \in L^\infty\Big([0,T); L^2(\Omega)\Big) \cap L^{r+1}(\Omega \times (0,T)), u_{tt}, v_{tt} \in L^\infty\Big([0,T), L^2(\Omega)\Big) \Big\}. \end{aligned} \tag{2.9}$$

Define the energy functional $E(t)$ associated with our system

$$E(t) = \frac{1}{2}\Big(\|u_t(t)\|_2^2 + \|v_t(t)\|_2^2\Big) + \frac{1}{2}\int_\Omega \Big(G\Big(|\nabla u|^2\Big) + G\Big(|\nabla v|^2\Big)\Big)dx - \int_\Omega F(u,v)dx. \tag{2.10}$$

A simple computation gives

$$\frac{dE(t)}{dt} = -\|u\|_{m+1}^{m+1} - \|v\|_{r+1}^{r+1} \le 0. \tag{2.11}$$

Our main result reads as follows.

Theorem 2.2. *Assume that* (2.3) *holds. Assume further that* $2(p+2) > \max\{2q+2, m+1, r+1\}$. *Then any solution of* (1.1) *with initial data satisfying*

$$\left(\int_\Omega \Big(G\Big(|\nabla u_0|^2\Big) + G\Big(|\nabla v_0|^2\Big)\Big)dx\right)^{1/2} > \alpha_1, \qquad E(0) < E_2, \tag{2.12}$$

cannot exist for all time, where the constant α_1 *and* E_2 *are defined in* (3.7).

3. Proof of Theorem 2.2

In this section, we deal with the blow-up of solutions of the system (1.1). Before we prove our main result, we need the following lemmas.

Lemma 3.1. *Let* $\Theta(t)$ *be a solution of the ordinary differential inequality*

$$\frac{d\Theta(t)}{dt} \ge C\Theta^{1+\varepsilon}(t), \quad t > 0, \tag{3.1}$$

where $\varepsilon > 0$. *If* $\Theta(0) > 0$, *then the solution ceases to exist for* $t \ge \Theta^{-\varepsilon}(0)C^{-1}\varepsilon^{-1}$.

Lemma 3.2. *Assume that* (2.3) *holds. Then there exists* $\eta > 0$ *such that for any* $(u, v) \in Z$, *one has*

$$\|u+v\|_{2(p+2)}^{2(p+2)} + 2\|uv\|_{p+2}^{p+2} \le \eta\left(\int_\Omega \Big(G\Big(|\nabla u|^2\Big) + G\Big(|\nabla v|^2\Big)\Big)dx\right)^{p+2}. \tag{3.2}$$

to Contents, page 1

Proof. By using Minkowski's inequality, we get

$$\|u+v\|_{2(p+2)}^2 \le 2\Big(\|u\|_{2(p+2)}^2 + \|v\|_{2(p+2)}^2\Big). \tag{3.3}$$

Also, Hölder's and Young's inequalities give us

$$\|uv\|_{p+2} \le \|u\|_{2(p+2)}\|v\|_{2(p+2)} \le \frac{1}{2}\Big(\|u\|_{2(p+2)}^2 + \|v\|_{2(p+2)}^2\Big). \tag{3.4}$$

If $b_1 > 0$, then we have

$$\int_\Omega \Big(G\big(|\nabla u|^2\big) + G\big(|\nabla v|^2\big)\Big)dx \ge c\Big(\|\nabla u\|_2^2 + \|\nabla v\|_2^2\Big). \tag{3.5}$$

If $b_1 = 0$, from $b_1 + b_2 > 0$, we have $b_2 > 0$. Since $W_0^{1,2(q+1)}(\Omega) \hookrightarrow H_0^1(\Omega)$, we have

$$\|\nabla u\|_2^2 + \|\nabla v\|_2^2 \le c_1\Big(\|\nabla u\|_{2(q+1)}^2 + \|\nabla v\|_{2(q+1)}^2\Big), \tag{3.6}$$

which implies that (3.5) still holds for $b_1 = 0$. Combining (3.3), (3.4) with (3.5) and the embedding $H_0^1(\Omega) \hookrightarrow L^{2(p+2)}(\Omega)$, we have (3.2). □

In order to prove our result and for the sake of simplicity, we take $a = b = 1$ and introduce the following:

$$\begin{gathered} B = \eta^{1/(2(p+2))}, \qquad \alpha_1 = B^{-(p+2/(p+1))}, \qquad E_1 = \left(\frac{1}{2} - \frac{1}{2(p+2)}\right)\alpha_1^2, \\ E_2 = \left(\frac{1}{2(q+1)} - \frac{1}{2(p+2)}\right)\alpha_1^2, \end{gathered} \tag{3.7}$$

where η is the optimal constant in (3.2). The following lemma will play an essential role in the proof of our main result, and it is similar to a lemma used first by Vitillaro [29].

Lemma 3.3. *Assume that* (2.3) *holds. Let* $(u, v) \in Z$ *be the solution of the system* (1.1). *Assume further that* $E(0) < E_1$ *and*

$$\left(\int_\Omega \Big(G\big(|\nabla u_0|^2\big) + G\big(|\nabla v_0|^2\big)\Big)dx\right)^{1/2} > \alpha_1. \tag{3.8}$$

Then there exists a constant $\alpha_2 > \alpha_1$ *such that*

$$\left(\int_\Omega \Big(G\big(|\nabla u|^2\big) + G\big(|\nabla v|^2\big)\Big)dx\right)^{1/2} \ge \alpha_2, \quad \text{for } t > 0, \tag{3.9}$$

$$\Big(\|u+v\|_{2(p+2)}^{2(p+2)} + 2\|uv\|_{p+2}^{p+2}\Big)^{1/(2(p+2))} \ge B\alpha_2, \quad \text{for } t > 0. \tag{3.10}$$

to Contents, page 1

Proof. We first note that, by (2.10), (3.2), and the definition of B, we have

$$\begin{aligned} E(t) &\geq \frac{1}{2}\int_\Omega\left(G\left(|\nabla u|^2\right)+G\left(|\nabla v|^2\right)\right)dx-\frac{1}{2(p+2)}\left(\|u+v\|_{2(p+2)}^{2(p+2)}+2\|uv\|_{p+2}^{p+2}\right)\\ &\geq \frac{1}{2}\int_\Omega\left(G\left(|\nabla u|^2\right)+G\left(|\nabla v|^2\right)\right)dx-\frac{B^{2(p+2)}}{2(p+2)}\left(\int_\Omega\left(G\left(|\nabla u|^2\right)+G\left(|\nabla v|^2\right)\right)dx\right)^{p+2}\\ &= \frac{1}{2}\alpha^2-\frac{B^{2(p+2)}}{2(p+2)}\alpha^{2(p+2)}, \end{aligned} \tag{3.11}$$

where $\alpha = (\int_\Omega(G(|\nabla u|^2)+G(|\nabla v|^2))dx)^{1/2}$. It is not hard to verify that g is increasing for $0<\alpha<\alpha_1$, decreasing for $\alpha>\alpha_1$, $g(\alpha)\to-\infty$ as $\alpha\to+\infty$, and

$$g(\alpha_1)=\frac{1}{2}\alpha_1^2-\frac{B^{2(p+2)}}{2(p+2)}\alpha_1^{2(p+2)}=E_1, \tag{3.12}$$

where α_1 is given in (3.7). Since $E(0)<E_1$, there exists $\alpha_2>\alpha_1$ such that $g(\alpha_2)=E(0)$.

Set $\alpha_0=(\int_\Omega(G(|\nabla u_0|^2)+G(|\nabla v_0|^2))dx)^{1/2}$. Then by (3.11) we get $g(\alpha_0)\leq E(0)=g(\alpha_2)$, which implies that $\alpha_0\geq\alpha_2$. Now, to establish (3.9), we suppose by contradiction that

$$\left(\int_\Omega\left(G\left(|\nabla u(t_0)|^2\right)+G\left(|\nabla v(t_0)|^2\right)\right)dx\right)^{1/2}<\alpha_2, \tag{3.13}$$

for some $t_0>0$. By the continuity of $\int_\Omega(G(|\nabla u|^2)+G(|\nabla v|^2))dx$, we can choose t_0 such that

$$\left(\int_\Omega\left(G\left(|\nabla u(t_0)|^2\right)+G\left(|\nabla v(t_0)|^2\right)\right)dx\right)^{1/2}>\alpha_1. \tag{3.14}$$

Again, the use of (3.11) leads to

$$E(t_0)\geq g\left(\left(\int_\Omega\left(G\left(|\nabla u(t_0)|^2\right)+G\left(|\nabla v(t_0)|^2\right)\right)dx\right)^{1/2}\right)>g(\alpha_2)=E(0). \tag{3.15}$$

This is impossible since $E(t)\leq E(0)$ for all $t\in[0,T)$. Hence (3.9) is established.

To prove (3.10), we make use of (2.10) to get

$$\frac{1}{2}\int_\Omega\left(G\left(|\nabla u|^2\right)+G\left(|\nabla v|^2\right)\right)dx\leq E(0)+\frac{1}{2(p+2)}\left(\|u+v\|_{2(p+2)}^{2(p+2)}+2\|uv\|_{p+2}^{p+2}\right). \tag{3.16}$$

to Contents, page 1

Consequently, (3.9) yields

$$\begin{aligned}\frac{1}{2(p+2)}\left(\|u+v\|_{2(p+2)}^{2(p+2)}+2\|uv\|_{p+2}^{p+2}\right) &\geq \frac{1}{2}\int_{\Omega}\left(G\left(|\nabla u|^2\right)+G\left(|\nabla v|^2\right)\right)dx-E(0)\\ &\geq \frac{1}{2}\alpha_2^2-E(0)\geq \frac{1}{2}\alpha_2^2-g(\alpha_2)=\frac{B^{2(p+2)}}{2(p+2)}\alpha_2^{2(p+2)}.\end{aligned}\tag{3.17}$$

Therefore, (3.17) and (3.7) yield the desired result. □

Proof of Theorem 2.2. We suppose that the solution exists for all time and we reach to a contradiction. Set

$$H(t)=E_2-E(t).\tag{3.18}$$

By using (2.10) and (3.18), we have

$$\begin{aligned}0<H(0)\leq H(t)=E_2&-\frac{1}{2}\left(\|u_t(t)\|_2^2+\|v_t(t)\|_2^2\right)-\frac{1}{2}\int_{\Omega}\left(G\left(|\nabla u|^2\right)+G\left(|\nabla v|^2\right)\right)dx\\ &+\frac{1}{2(p+2)}\left(\|u+v\|_{2(p+2)}^{2(p+2)}+2\|uv\|_{p+2}^{p+2}\right).\end{aligned}\tag{3.19}$$

From (3.9), we have

$$\begin{aligned}E_2&-\frac{1}{2}\left(\|u_t(t)\|_2^2+\|v_t(t)\|_2^2\right)-\frac{1}{2}\int_{\Omega}\left(G\left(|\nabla u|^2\right)+G\left(|\nabla v|^2\right)\right)dx\\ &\leq E_2-\frac{1}{2}\alpha_1^2\leq E_1-\frac{1}{2}\alpha_1^2=-\frac{1}{2(p+2)}\alpha_1^2<0,\quad \forall t\geq 0.\end{aligned}\tag{3.20}$$

Hence, by the above inequality and (2.6), we have

$$0<H(0)\leq H(t)\leq \frac{1}{2(p+2)}\left(\|u+v\|_{2(p+2)}^{2(p+2)}+2\|uv\|_{p+2}^{p+2}\right),\tag{3.21}$$

$$\leq \frac{c_1}{2(p+2)}\left(\|u\|_{2(p+2)}^{2(p+2)}+\|v\|_{2(p+2)}^{2(p+2)}\right).\tag{3.22}$$

We then define

$$\Theta(t)=H^{1-\delta}(t)+\epsilon\int_{\Omega}(uu_t+vv_t)dx,\tag{3.23}$$

where ϵ small enough is to be chosen later and

$$0<\delta\leq \min\left\{\frac{p+1}{2(p+2)},\frac{2(p+2)-(m+1)}{2m(p+2)},\frac{2(p+2)-(r+1)}{2r(p+2)}\right\}.\tag{3.24}$$

to Contents, page 1

Our goal is to show that $\Theta(t)$ satisfies the differential inequality (3.1) which leads to a blow-up in finite time. By taking a derivative of (3.23), we get

$$\begin{aligned}\Theta'(t) &= (1-\delta)H^{-\delta}(t)H'(t) + \epsilon\left(\|u_t\|_2^2 + \|v_t\|_2^2\right) - \epsilon\int_\Omega\left(g\left(|\nabla u|^2\right)|\nabla u|^2 + g\left(|\nabla v|^2\right)|\nabla v|^2\right)dx \\ &\quad - \epsilon\int_\Omega\left(|u_t|^{m-1}u_t u + |v_t|^{r-1}v_t v\right)dx + \epsilon\int_\Omega\left(uf_1(u,v) + vf_2(u,v)\right)dx \\ &= (1-\delta)H^{-\delta}(t)H'(t) + \epsilon\left(\|u_t\|_2^2 + \|v_t\|_2^2\right) - b_1\epsilon\left(\|\nabla u\|_2^2 + \|\nabla v\|_2^2\right) - \epsilon b_2\|\nabla u\|_{2(q+2)}^{2(q+2)} \\ &\quad - \epsilon b_2\|\nabla v\|_{2(q+2)}^{2(q+2)} - \epsilon\int_\Omega\left(|u_t|^{m-1}u_t u + |v_t|^{r-1}v_t v\right)dx + \epsilon\left(\|u+v\|_{2(p+2)}^{2(p+2)} + 2\|uv\|_{p+2}^{p+2}\right).\end{aligned} \tag{3.25}$$

From the definition of $H(t)$, it follows that

$$\begin{aligned}-b_2\|\nabla u\|_{2(q+2)}^{2(q+2)} - b_2\|\nabla v\|_{2(q+2)}^{2(q+2)} &= 2(q+1)H(t) - 2(q+1)E_2 + (q+1)\left(\|u_t\|_2^2 + \|v_t\|_2^2\right) \\ &\quad + (q+1)b_1\left(\|\nabla u\|_2^2 + \|\nabla v\|_2^2\right) - 2(q+1)\int_\Omega F(u,v)dx,\end{aligned} \tag{3.26}$$

which together with (3.25) gives

$$\begin{aligned}\Theta'(t) &= (1-\delta)H^{-\delta}(t)H'(t) + \epsilon(q+2)\left(\|u_t\|_2^2 + \|v_t\|_2^2\right) + b_1 q\epsilon\left(\|\nabla u\|_2^2 + \|\nabla v\|_2^2\right) \\ &\quad - \epsilon\int_\Omega\left(|u_t|^{m-1}u_t u + |v_t|^{r-1}v_t v\right)dx + \epsilon\left(1 - \frac{q+1}{p+2}\right)\left(\|u+v\|_{2(p+2)}^{2(p+2)} + 2\|uv\|_{p+2}^{p+2}\right) \\ &\quad + 2(q+1)H(t) - 2(q+1)E_2.\end{aligned} \tag{3.27}$$

Then, using (3.10), we obtain

$$\begin{aligned}\Theta'(t) &\geq (1-\delta)H^{-\delta}(t)H'(t) + \epsilon(q+2)\left(\|u_t\|_2^2 + \|v_t\|_2^2\right) + b_1 q\epsilon\left(\|\nabla u\|_2^2 + \|\nabla v\|_2^2\right) + 2(q+1)H(t) \\ &\quad + \epsilon\overline{c}\left(\|u+v\|_{2(p+2)}^{2(p+2)} + 2\|uv\|_{p+2}^{p+2}\right) - \epsilon\int_\Omega\left(|u_t|^{m-1}u_t u + |v_t|^{r-1}v_t v\right)dx,\end{aligned} \tag{3.28}$$

where $\overline{c} = 1-(q+1)/(p+2)-2(q+1)E_2(B\alpha_2)^{-2(p+2)}$. It is clear that $\overline{c} > 0$, since $\alpha_2 > B^{-(p+2)/(p+1)}$. We now exploit Young's inequality to estimate the last two terms on the right side of (3.28)

$$\begin{aligned}\left|\int_\Omega |u_t|^{m-1}u_t u dx\right| &\leq \frac{\eta_1^{m+1}}{m+1}\|u\|_{m+1}^{m+1} + \frac{m\eta_1^{-((m+1)/m)}}{m+1}\|u_t\|_{m+1}^{m+1}, \\ \left|\int_\Omega |v_t|^{r-1}v_t v dx\right| &\leq \frac{\eta_2^{r+1}}{r+1}\|v\|_{r+1}^{r+1} + \frac{r\eta_2^{-((r+1)/r)}}{r+1}\|v_t\|_{r+1}^{r+1},\end{aligned} \tag{3.29}$$

to Contents, page 1

where η_1, η_2 are parameters depending on the time t and specified later. Inserting the last two estimates into (3.28), we have

$$\Theta'(t) \geq (1-\delta)H^{-\delta}(t)H'(t) + \epsilon(q+2)\left(\|u_t\|_2^2 + \|v_t\|_2^2\right) + b_1 q\epsilon\left(\|\nabla u\|_2^2 + \|\nabla v\|_2^2\right) + 2(q+1)H(t)$$
$$+ \epsilon\bar{c}\left(\|u+v\|_{2(p+2)}^{2(p+2)} + 2\|uv\|_{p+2}^{p+2}\right) - \epsilon\frac{\eta_1^{m+1}}{m+1}\|u\|_{m+1}^{m+1} - \epsilon\frac{m\eta_1^{-((m+1)/m)}}{m+1}\|u_t\|_{m+1}^{m+1}$$
$$- \epsilon\frac{\eta_2^{r+1}}{r+1}\|v\|_{r+1}^{r+1} - \epsilon\frac{r\eta_2^{-((r+1)/r)}}{r+1}\|v_t\|_{r+1}^{r+1}. \tag{3.30}$$

By choosing η_1 and η_2 such that

$$\eta_1^{-(m+1)/m} = M_1 H^{-\delta}(t), \qquad \eta_2^{-(r+1)/r} = M_2 H^{-\delta}(t), \tag{3.31}$$

where M_1 and M_2 are constants to be fixed later. Thus, by using (2.6) and (3.31), inequality (3.31) then takes the form

$$\Theta'(t) \geq ((1-\delta) - M\epsilon)H^{-\delta}(t)H'(t) + \epsilon(q+2)\left(\|u_t\|_2^2 + \|v_t\|_2^2\right) + b_1 q\epsilon\left(\|\nabla u\|_2^2 + \|\nabla v\|_2^2\right)$$
$$+ 2(q+1)H(t) + \epsilon c_2\left(\|u\|_{2(p+2)}^{2(p+2)} + 2\|v\|_{2(p+2)}^{2(p+2)}\right) - \epsilon M_1^{-m} H^{\delta m}(t)\|u\|_{m+1}^{m+1}$$
$$- \epsilon M_2^{-r} H^{\delta r}(t)\|v\|_{r+1}^{r+1}, \tag{3.32}$$

where $M = m/(m+1)M_1 + r/(r+1)M_2$ and c_2 is a positive constant.

Since $2(p+2) > \max\{m+1, r+1\}$, taking into account (2.6) and (3.21), then we have

$$\begin{aligned} H^{\delta m}(t)\|u\|_{m+1}^{m+1} &\leq c_3\left(\|u\|_{2(p+2)}^{2\delta m(p+2)+(m+1)} + \|v\|_{2(p+2)}^{2\delta m(p+2)}\|u\|_{m+1}^{m+1}\right), \\ H^{\delta r}(t)\|v\|_{r+1}^{r+1} &\leq c_4\left(\|v\|_{2(p+2)}^{2\delta r(p+2)+(r+1)} + \|u\|_{2(p+2)}^{2\delta r(p+2)}\|v\|_{r+1}^{r+1}\right), \end{aligned} \tag{3.33}$$

for some positive constants c_3 and c_4. By using (3.24) and the algebraic inequality

$$z^\nu \leq z + 1 \leq \left(1 + \frac{1}{a}\right)(z + a), \quad \forall z \geq 0,\ 0 < \nu \leq 1,\ a \geq 0, \tag{3.34}$$

we have

$$\|u\|_{2(p+2)}^{2\delta m(p+2)+(m+1)} \leq d\left(\|u\|_{2(p+2)}^{2(p+2)} + H(0)\right) \leq d\left(\|u\|_{2(p+2)}^{2(p+2)} + H(t)\right), \quad \forall t \geq 0, \tag{3.35}$$

to Contents, page 1

where $d = 1 + 1/H(0)$. Similarly,

$$\|v\|_{2(p+2)}^{2\delta r(p+2)+(r+1)} \le d\left(\|v\|_{2(p+2)}^{2(p+2)} + H(t)\right), \quad \forall t \ge 0. \tag{3.36}$$

Also, since

$$(X+Y)^s \le C(X^s + Y^s), \quad X, Y \ge 0, \ s > 0, \tag{3.37}$$

by using (3.24) and (3.34), we conclude that

$$\begin{aligned} \|v\|_{2(p+2)}^{2\delta m(p+2)} \|u\|_{m+1}^{m+1} &\le C\left(\|v\|_{2(p+2)}^{2(p+2)} + \|u\|_{(m+1)}^{2(p+2)}\right) \le C\left(\|v\|_{2(p+2)}^{2(p+2)} + \|u\|_{2(p+2)}^{2(p+2)}\right), \\ \|u\|_{2(p+2)}^{2\delta r(p+2)} \|v\|_{r+1}^{r+1} &\le C\left(\|u\|_{2(p+2)}^{2(p+2)} + \|v\|_{(r+1)}^{2(p+2)}\right) \le C\left(\|u\|_{2(p+2)}^{2(p+2)} + \|v\|_{2(p+2)}^{2(p+2)}\right), \end{aligned} \tag{3.38}$$

where C is a generic positive constant. Taking into account (3.33)–(3.38), estimate (3.32) takes the form

$$\begin{aligned} \Theta'(t) \ge{}& ((1-\delta) - M\epsilon) H^{-\delta}(t) H'(t) + \epsilon(q+2)\left(\|u_t\|_2^2 + \|v_t\|_2^2\right) \\ &+ \epsilon\left(2(q+1) - C_1 M_1^{-m} - C_1 M_2^{-r}\right) H(t) \\ &+ \epsilon\left(c_2 - C_2 M_1^{-m} - C_2 M_2^{-r}\right)\left(\|u\|_{2(p+2)}^{2(p+2)} + \|v\|_{2(p+2)}^{2(p+2)}\right), \end{aligned} \tag{3.39}$$

where $C_1 = \max\{c_3 d + C, c_4 d + C\}$, $C_2 = \max\{c_3 d, c_4 d\}$. At this point, and for large values of M_1 and M_2, we can find positive constants κ_1 and κ_2 such that (3.39) becomes

$$\begin{aligned} \Theta'(t) \ge{}& ((1-\delta) - M\epsilon) H^{-\delta}(t) H'(t) + \epsilon(q+2)\left(\|u_t\|_2^2 + \|v_t\|_2^2\right) \\ &+ \epsilon\kappa_1 H(t) + \epsilon\kappa_2\left(\|u\|_{2(p+2)}^{2(p+2)} + \|v\|_{2(p+2)}^{2(p+2)}\right). \end{aligned} \tag{3.40}$$

Once M_1 and M_2 are fixed, we pick ϵ small enough so that $(1-\delta) - M\epsilon \ge 0$ and

$$\Theta(0) = H^{1-\delta}(0) + \epsilon \int_\Omega (u_0 u_1 + v_0 v_1)\, dx > 0. \tag{3.41}$$

Since $H'(t) \ge 0$, there exists $\Lambda > 0$ such that (3.40) becomes

$$\Theta'(t) \ge \epsilon\Lambda\left(H(t) + \|u_t\|_2^2 + \|v_t\|_2^2 + \|u\|_{2(p+2)}^{2(p+2)} + \|v\|_{2(p+2)}^{2(p+2)}\right). \tag{3.42}$$

Then, we have

$$\Theta(t) \ge \Theta(0), \quad \forall t \ge 0. \tag{3.43}$$

to Contents, page 1

Next, we have by Hölder's and Young's inequalities

$$\left(\int_{\Omega} uu_t dx + \int_{\Omega} vv_t dx\right)^{1/(1-\delta)} \leq C\left(\|u\|_{2(p+2)}^{\tau/(1-\delta)} + \|u_t\|_2^{s/(1-\delta)} + \|v\|_{2(p+2)}^{\tau/(1-\delta)} + \|v_t\|_2^{s/(1-\delta)}\right), \tag{3.44}$$

for $1/\tau + 1/s = 1$. We take $s = 2(1-\delta)$, to get $\tau/(1-\delta) = 2/(1-2\delta)$. Here and in the sequel, C denotes a positive constant which may change from line to line. By using (3.24) and (3.34), we have

$$\|u\|_{2(p+2)}^{2/(1-2\delta)} \leq d\left(\|u\|_{2(p+2)}^{2(p+2)} + H(t)\right), \quad \|v\|_{2(p+2)}^{2/(1-2\delta)} \leq d\left(\|v\|_{2(p+2)}^{2(p+2)} + H(t)\right), \quad \forall t \geq 0. \tag{3.45}$$

Therefore, (3.44) becomes

$$\left(\int_{\Omega} uu_t dx + \int_{\Omega} vv_t dx\right)^{1/(1-\delta)} \leq C\left(\|u\|_{2(p+2)}^{2(p+2)} + \|v\|_{2(p+2)}^{2(p+2)} + \|u_t\|_2^2 + \|v_t\|_2^2\right). \tag{3.46}$$

Note that

$$\begin{aligned}
\Theta^{1/(1-\delta)}(t) &= \left(H^{1-\delta}(t) + \epsilon\int_{\Omega}(uu_t + vv_t)dx\right)^{1/(1-\delta)} \\
&\leq C\left(H(t) + \left|\int_{\Omega} uu_t dx + \int_{\Omega} vv_t dx\right|^{1/(1-\delta)}\right) \\
&\leq C\left(H(t) + \|u\|_{2(p+2)}^{2(p+2)} + \|v\|_{2(p+2)}^{2(p+2)} + \|u_t\|_2^2 + \|v_t\|_2^2\right).
\end{aligned} \tag{3.47}$$

Combining (3.42) with (3.47), we have

$$\Theta(t) \geq C\Theta^{1/(1-\delta)}(t), \quad \forall t \geq 0. \tag{3.48}$$

A simple application of Lemma 3.1 gives the desired result. □

Acknowledgments

The authors are indebted to the referee for giving some important suggestions which improved the presentations of this paper. This work is supported in part by a China NSF Grant no. 10871097, Qing Lan Project of Jiangsu Province, the Foundation for Young Talents in College of Anhui Province Grant no. 2011SQRL115 and Program sponsored for scientific innovation research of college graduate in Jangsu province.

to Contents, page 1

References

[1] A. Haraux and E. Zuazua, "Decay estimates for some semilinear damped hyperbolic problems," *Archive for Rational Mechanics and Analysis*, vol. 100, no. 2, pp. 191–206, 1988.
[2] J. M. Ball, "Remarks on blow-up and nonexistence theorems for nonlinear evolution equations," *The Quarterly Journal of Mathematics. Oxford. Second Series*, vol. 28, no. 112, pp. 473–486, 1977.
[3] V. K. Kalantarov, "The occurrence of collapse for quasilinear equations of parabolic and hyperbolic types," *Journal of Soviet Mathematics*, vol. 10, no. 1, pp. 53–70, 1978.
[4] H. A. Levine, "Instability and nonexistence of global solutions to nonlinear wave equations of the form $Pu_{tt} = -Au + \mathcal{F}(u)$," *Transactions of the American Mathematical Society*, vol. 192, pp. 1–21, 1974.
[5] H. A. Levine, "Some additional remarks on the nonexistence of global solutions to nonlinear wave equations," *SIAM Journal on Mathematical Analysis*, vol. 5, pp. 138–146, 1974.
[6] V. Georgiev and G. Todorova, "Existence of a solution of the wave equation with nonlinear damping and source terms," *Journal of Differential Equations*, vol. 109, no. 2, pp. 295–308, 1994.
[7] S. A. Messaoudi, "Blow up in a nonlinearly damped wave equation," *Mathematische Nachrichten*, vol. 231, pp. 105–111, 2001.
[8] S. A. Messaoudi, "Blow-up of positive-initial-energy solutions of a nonlinear viscoelastic hyperbolic equation," *Journal of Mathematical Analysis and Applications*, vol. 320, no. 2, pp. 902–915, 2006.
[9] G. Andrews, "On the existence of solutions to the equation $u_{tt} = u_{xxt} + \sigma(u_x)_x$," *Journal of Differential Equations*, vol. 35, no. 2, pp. 200–231, 1980.
[10] G. Andrews and J. M. Ball, "Asymptotic behaviour and changes of phase in one-dimensional nonlinear viscoelasticity," *Journal of Differential Equations*, vol. 44, no. 2, pp. 306–341, 1982.
[11] J. Clements, "Existence theorems for a quasilinear evolution equation," *SIAM Journal on Applied Mathematics*, vol. 26, pp. 745–752, 1974.
[12] S. Kawashima and Y. Shibata, "Global existence and exponential stability of small solutions to nonlinear viscoelasticity," *Communications in Mathematical Physics*, vol. 148, no. 1, pp. 189–208, 1992.
[13] D. D Ang and A. P. N. Dinh, "Strong solutions of a quasilinear wave equation with nonlinear damping," *SIAM Journal on Mathematical Analysis*, vol. 19, no. 2, pp. 337–347, 1988.
[14] W. N. Findley, J. S. Lai, and K. Onaran, *Creep and Relaxation of Nonlinear Viscoelastic Materials*, North-Holland, Amsterdam, The Netherlands, 1976.
[15] J. K. Knowles, "On finite anti-plane shear for incompressible elastic materials," *Australian Mathematical Society. Journal. Series B*, vol. 19, no. 4, pp. 400–415, 1975/76.
[16] V. P. Maslov and P. P. Mosolov, *Nonlinear Wave Equations Perturbed by Viscous Terms*, vol. 31 of *de Gruyter Expositions in Mathematics*, Walter De Gruyter, Berlin, Germany, 2000.
[17] J. M. Greenberg, R. C. MacCamy, and V. J. Mizel, "On the existence, uniqueness, and stability of solutions of the equation $\sigma'(u_x)u_{xx} + \lambda u_{xxt} = \rho_0 u_{tt}$," vol. 17, pp. 707–728, 1967/1968.
[18] Y. Yamada, "Some remarks on the equation $Y_{tt} - \sigma(Y_x)Y_{xx} - Y_{xxt} = f$," *Osaka Journal of Mathematics*, vol. 17, no. 2, pp. 303–323, 1980.
[19] Z. Yang and G. Chen, "Global existence of solutions for quasi-linear wave equations with viscous damping," *Journal of Mathematical Analysis and Applications*, vol. 285, no. 2, pp. 604–618, 2003.
[20] G. Chen, H. Yue, and S. Wang, "The initial boundary value problem for quasi-linear wave equation with viscous damping," *Journal of Mathematical Analysis and Applications*, vol. 331, no. 2, pp. 823–839, 2007.
[21] J. Hao, Y. Zhang, and S. Li, "Global existence and blow-up phenomena for a nonlinear wave equation," *Nonlinear Analysis*, vol. 71, no. 10, pp. 4823–4832, 2009.
[22] S. A. Messaoudi and B. Said Houari, "Global non-existence of solutions of a class of wave equations with non-linear damping and source terms," *Mathematical Methods in the Applied Sciences*, vol. 27, no. 14, pp. 1687–1696, 2004.
[23] I. E. Segal, "Nonlinear partial differential equations in quantum field theory," in *Proc. Sympos. Appl. Math., Vol. XVII*, pp. 210–226, American Mathematical Society, Providence, RI, USA, 1965.
[24] K. Jorgens, *Nonlinear Wave Equations*, University of Colorado, Department of Mathematics, 1970.
[25] K. Agre and M. A. Rammaha, "Systems of nonlinear wave equations with damping and source terms," *Differential and Integral Equations*, vol. 19, no. 11, pp. 1235–1270, 2006.
[26] B. Said-Houari, "Global nonexistence of positive initial-energy solutions of a system of nonlinear wave equations with damping and source terms," *Differential and Integral Equations*, vol. 23, no. 1-2, pp. 79–92, 2010.
[27] J. Wu, S. Li, and S. Chai, "Existence and nonexistence of a global solution for coupled nonlinear wave equations with damping and source," *Nonlinear Analysis*, vol. 72, no. 11, pp. 3969–3975, 2010.

to Contents, page 1

[28] L. E. Payne and D. H. Sattinger, “Saddle points and instability of nonlinear hyperbolic equations,” *Israel Journal of Mathematics*, vol. 22, no. 3-4, pp. 273–303, 1975.
[29] E. Vitillaro, “Global nonexistence theorems for a class of evolution equations with dissipation,” *Archive for Rational Mechanics and Analysis*, vol. 149, no. 2, pp. 155–182, 1999.
[30] S. A. Messaoudi and B. Said-Houari, “Global nonexistence of positive initial-energy solutions of a system of nonlinear viscoelastic wave equations with damping and source terms,” *Journal of Mathematical Analysis and Applications*, vol. 365, no. 1, pp. 277–287, 2010.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 970978, 22 pages
doi:10.1155/2011/970978

Research Article

H_∞ Estimation for a Class of Lipschitz Nonlinear Discrete-Time Systems with Time Delay

Huihong Zhao, Chenghui Zhang, Guangchen Wang, and Guojing Xing

School of Control Science and Engineering, Shandong University, 17923 Jingshi Road, Jinan 250061, China

Correspondence should be addressed to Chenghui Zhang, zchui@sdu.edu.cn

Received 27 December 2010; Accepted 18 May 2011

Academic Editor: Elena Braverman

The issue of H_∞ estimation for a class of Lipschitz nonlinear discrete-time systems with time delay and disturbance input is addressed. First, through integrating the H_∞ filtering performance index with the Lipschitz conditions of the nonlinearity, the design of robust estimator is formulated as a positive minimum problem of indefinite quadratic form. Then, by introducing the Krein space model and applying innovation analysis approach, the minimum of the indefinite quadratic form is obtained in terms of innovation sequence. Finally, through guaranteeing the positivity of the minimum, a sufficient condition for the existence of the H_∞ estimator is proposed and the estimator is derived in terms of Riccati-like difference equations. The proposed algorithm is proved to be effective by a numerical example.

1. Introduction

In control field, nonlinear estimation is considered to be an important task which is also of great challenge, and it has been a very active area of research for decades [1–7]. Many kinds of methods on estimator design have been proposed for different types of nonlinear dynamical systems. Generally speaking, there are three approaches widely adopted for nonlinear estimation. In the first one, by using an extended (nonexact) linearization of the nonlinear systems, the estimator is designed by employing classical linear observer techniques [1]. The second approach, based on a nonlinear state coordinate transformation which renders the dynamics driven by nonlinear output injection and the output linear on the new coordinates, uses the quasilinear approaches to design the nonlinear estimator [2–4]. In the last one, methods are developed to design nonlinear estimators for systems which consist of an observable linear part and a locally or globally Lipschitz nonlinear part [5–7]. In this paper, the problem of H_∞ estimator design is investigated for a class of Lipschitz nonlinear discrete-time systems with time delay and disturbance input.

to Contents, page 1

In practice, most nonlinearities can be regarded as Lipschitz, at least locally when they are studied in a given neighborhood [6]. For example, trigonometric nonlinearities occurring in many robotic problems, non-linear softening spring models frequently used in mechanical systems, nonlinearities which are square or cubic in nature, and so forth. Thus, in recent years, increasing attention has been paid to estimator design for Lipschitz nonlinear systems [8–19]. For the purpose of designing this class of nonlinear estimator, a number of approaches have been developed, such as sliding mode observers [8, 9], H_∞ optimization techniques [10–13], adaptive observers [14, 15], high-gain observers [16], loop transfer recovery observers [17], proportional integral observers [18], and integral quadratic constraints approach [19]. All of the above results are obtained in the assumption that the Lipschitz nonlinear systems are delay free. However, time delay is an inherent characteristic of many physical systems, and it can result in instability and poor performances if it is ignored. The estimator design for time-delay Lipschitz nonlinear systems has become a substantial need. Unfortunately, compared with estimator design for delay-free Lipschitz nonlinear systems, less research has been carried out on the time-delay case. In [20], the linear matrix inequality-(LMI-) based full-order and reduced-order robust H_∞ observers are proposed for a class of Lipschitz nonlinear discrete-time systems with time delay. In [21], by using Lyapunov stability theory and LMI techniques, a delay-dependent approach to the H_∞ and L_2−L_∞ filtering is proposed for a class of uncertain Lipschitz nonlinear time-delay systems. In [22], by guaranteeing the asymptotic stability of the error dynamics, the robust observer is presented for a class of uncertain discrete-time Lipschitz nonlinear state delayed systems; In [23], based on the sliding mode techniques, a discontinuous observer is designed for a class of Lipschitz nonlinear systems with uncertainty. In [24], an LMI-based convex optimization approach to observer design is developed for both constant-delay and time-varying delay Lipschitz nonlinear systems.

In this paper, the H_∞ estimation problem is studied for a class of Lipschitz nonlinear discrete time-delay systems with disturbance input. Inspired by the recent study on H_∞ fault detection for linear discrete time-delay systems in [25], a recursive Kalman-like algorithm in an indefinite metric space, named the Krein space [26], will be developed to the design of H_∞ estimator for time-delay Lipschitz nonlinear systems. Unlike [20], the delay-free nonlinearities and the delayed nonlinearities in the presented systems are decoupling. For the case presented in [20], the H_∞ observer design problem, utilizing the technical line of this paper, can be solved by transforming it into a delay-free system through state augmentation. Indeed, the state augmentation results in a higher system dimension and, thus, a much more expensive computational cost. Therefore, this paper based on the presented time-delay Lipschitz nonlinear systems, focuses on the robust estimator design without state augmentation by employing innovation analysis approach in the Krein space. The major contribution of this paper can be summarized as follows: (i) it extends the Krein space linear estimation methodology [26] to the state estimation of the time-delay Lipschitz nonlinear systems and (ii) it develops a recursive Kalman-like robust estimator for time-delay Lipschitz nonlinear systems without state augmentation.

The remainder of this paper is arranged as follows. In Section 2, the interest system, the Lipschitz conditions, and the H_∞ estimation problem are introduced. In Section 3, a partially equivalent Krein space problem is constructed, the H_∞ estimator is obtained by computed Riccati-like difference equations, and sufficient existence condition is derived in terms of matrix inequalities. An example is given to show the effect of the proposed algorithm in Section 4. Finally, some concluding remarks are made in Section 5.

In the sequel, the following notation will be used: elements in the Krein space will be denoted by **boldface** letters, and elements in the Euclidean space of complex numbers

to Contents, page 1

will be denoted by normal letters; $\mathbb{R}^n$ denotes the real n-dimensional Euclidean space; $\|\cdot\|$ denotes the Euclidean norm; $\theta(k) \in l_2[0,N]$ means $\sum_{k=0}^{N}(\theta^T(k)\theta(k)) < \infty$; the superscripts "$-1$" and "$T$" stand for the inverse and transpose of a matrix, resp.; I is the identity matrix with appropriate dimensions; For a real matrix, $P > 0$ ($P < 0$, resp.) means that P is symmetric and positive (negative, resp.) definite; $\langle *, * \rangle$ denotes the inner product in the Krein space; diag$\{\cdots\}$ denotes a block-diagonal matrix; $\mathcal{L}\{\cdots\}$ denotes the linear space spanned by sequence $\{\cdots\}$.

2. System Model and Problem Formulation

Consider a class of nonlinear systems described by the following equations:

$$\begin{aligned} x(k+1) &= Ax(k) + A_d x(k_d) + f(k, Fx(k), u(k)) \\ &\quad + h(k, Hx(k_d), u(k)) + Bw(k), \\ y(k) &= Cx(k) + v(k), \\ z(k) &= Lx(k), \end{aligned} \tag{2.1}$$

where $k_d = k - d$, and the positive integer d denotes the known state delay; $x(k) \in \mathbb{R}^n$ is the state, $u(k) \in \mathbb{R}^p$ is the measurable information, $w(k) \in \mathbb{R}^q$ and $v(k) \in \mathbb{R}^m$ are the disturbance input belonging to $l_2[0,N]$, $y(k) \in \mathbb{R}^m$ is the measurement output, and $z(k) \in \mathbb{R}^r$ is the signal to be estimated; the initial condition $x_0(s)$ $(s = -d, -d+1, \ldots, 0)$ is unknown; the matrices $A \in \mathbb{R}^{n\times n}$, $A_d \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times q}$, $C \in \mathbb{R}^{m\times n}$ and $L \in \mathbb{R}^{r\times n}$, are real and known constant matrices.

In addition, $f(k, Fx(k), u(k))$ and $h(k, Hx(k_d), u(k))$ are assumed to satisfy the following Lipschitz conditions:

$$\begin{aligned} \|f(k, Fx(k), u(k)) - f(k, F\check{x}(k), u(k))\| &\le \alpha \|F(x(k) - \check{x}(k))\|, \\ \|h(k, Hx(k_d), u(k)) - h(k, H\check{x}(k_d), u(k))\| &\le \beta \|H(x(k_d) - \check{x}(k_d))\|, \end{aligned} \tag{2.2}$$

for all $k \in \{0, 1, \ldots, N\}$, $u(k) \in \mathbb{R}^p$ and $x(k), \check{x}(k), x(k_d), \check{x}(k_d) \in \mathbb{R}^n$. where $\alpha > 0$ and $\beta > 0$ are known Lipschitz constants, and F, H are real matrix with appropriate dimension.

The H_∞ estimation problem under investigation is stated as follows. Given the desired noise attenuation level $\gamma > 0$ and the observation $\{y(j)\}_{j=0}^{k}$, find an estimate $\check{z}(k \mid k)$ of the signal $z(k)$, if it exists, such that the following inequality is satisfied:

$$\sup_{(x_0, w, v) \ne 0} \frac{\sum_{k=0}^{N} \|\check{z}(k \mid k) - z(k)\|^2}{\sum_{k=-d}^{0} \|x_0(k)\|^2_{\Pi^{-1}(k)} + \sum_{k=0}^{N} \|w(k)\|^2 + \sum_{k=0}^{N} \|v(k)\|^2} < \gamma^2, \tag{2.3}$$

where $\Pi(k)$ $(k = -d, -d+1, \ldots, 0)$ is a given positive definite matrix function which reflects the relative uncertainty of the initial state $x_0(k)$ $(k = -d, -d+1, \ldots, 0)$ to the input and measurement noises.

Remark 2.1. For the sake of simplicity, the initial state estimate $\hat{x}_0(k)$ $(k = -d, -d+1, \ldots, 0)$ is assumed to be zero in inequality (2.3).

to Contents, page 1

Remark 2.2. Although the system given in [20] is different from the one given in this paper, the problem mentioned in [20] can also be solved by using the presented approach. The resolvent first converts the system given in [20] into a delay-free one by using the classical system augmentation approach, and then designs estimator by employing the similar but easier technical line with our paper.

3. Main Results

In this section, the Krein space-based approach is proposed to design the H_∞ estimator for Lipschitz nonlinear systems. To begin with, the H_∞ estimation problem (2.3) and the Lipschitz conditions (2.2) are combined in an indefinite quadratic form, and the nonlinearities are assumed to be obtained by $\{y(i)\}_{i=0}^{k}$ at the time step k. Then, an equivalent Krein space problem is constructed by introducing an imaginary Krein space stochastic system. Finally, based on projection formula and innovation analysis approach in the Krein space, the recursive estimator is derived.

3.1. Construct a Partially Equivalent Krein Space Problem

It is proved in this subsection that the H_∞ estimation problem can be reduced to a positive minimum problem of indefinite quadratic form, and the minimum can be obtained by using the Krein space-based approach.

Since the denominator of the left side of (2.3) is positive, the inequality (2.3) is equivalent to

$$\underbrace{\sum_{k=-d}^{0} \|x_0(k)\|^2_{\Pi^{-1}(k)} + \sum_{k=0}^{N} \|w(k)\|^2 + \sum_{k=0}^{N} \|v(k)\|^2 - \gamma^{-2} \sum_{k=0}^{N} \|v_z(k)\|^2}_{\triangleq J_N^*} > 0, \quad \forall (x_0, w, v) \neq 0, \tag{3.1}$$

where $v_z(k) = \check{z}(k \mid k) - z(k)$.

Moreover, we denote

$$\begin{aligned} z_f(k) &= Fx(k), & \check{z}_f(k \mid k) &= F\check{x}(k \mid k), \\ z_h(k_d) &= Hx(k_d), & \check{z}_h(k_d \mid k) &= H\check{x}(k_d \mid k), \end{aligned} \tag{3.2}$$

where $\check{z}_f(k \mid k)$ and $\check{z}_h(k_d \mid k)$ denote the optimal estimation of $z_f(k)$ and $z_h(k_d)$ based on the observation $\{y(j)\}_{j=0}^{k}$, respectively. And, let

$$\begin{aligned} w_f(k) &= f\left(k, z_f(k), u(k)\right) - f\left(k, \check{z}_f(k \mid k), u(k)\right), \\ w_h(k_d) &= h(k, z_h(k_d), u(k)) - h(k, \check{z}_h(k_d \mid k), u(k)), \\ v_{z_f}(k) &= \check{z}_f(k \mid k) - z_f(k), \\ v_{z_h}(k_d) &= \check{z}_h(k_d \mid k) - z_h(k_d). \end{aligned} \tag{3.3}$$

to Contents, page 1

From the Lipschitz conditions (2.2), we derive that

$$\underbrace{J_N^* + \sum_{k=0}^{N}\|w_f(k)\|^2 + \sum_{k=0}^{N}\|w_h(k_d)\|^2 - \alpha^2\sum_{k=0}^{N}\left\|v_{z_f}(k)\right\|^2 - \beta^2\sum_{k=0}^{N}\|v_{z_h}(k_d)\|^2}_{\triangleq J_N} \le J_N^*. \tag{3.4}$$

Note that the left side of (3.1) and (3.4), J_N, can be recast into the form

$$\begin{aligned} J_N = & \sum_{k=-d}^{0}\|x_0(k)\|^2_{\overline{\Pi}^{-1}(k)} + \sum_{k=0}^{N}\|\overline{w}(k)\|^2 + \sum_{k=0}^{N}\|v(k)\|^2 \\ & - \gamma^{-2}\sum_{k=0}^{N}\|v_z(k)\|^2 - \alpha^2\sum_{k=0}^{N}\left\|v_{z_f}(k)\right\|^2 - \beta^2\sum_{k=d}^{N}\|v_{z_h}(k_d)\|^2, \end{aligned} \tag{3.5}$$

where

$$\begin{aligned} \overline{\Pi}(k) &= \begin{cases} \left(\Pi^{-1}(k) - \beta^2 H^T H\right)^{-1}, & k = -d, \ldots, -1, \\ \Pi(k), & k = 0, \end{cases} \\ \overline{w}(k) &= \left[w^T(k) \;\; w_f^T(k) \;\; w_h^T(k_d)\right]^T. \end{aligned} \tag{3.6}$$

Since $J_N \le J_N^*$, it is natural to see that if $J_N > 0$ then the H_∞ estimation problem (2.3) is satisfied, that is, $J_N^* > 0$. Hence, the H_∞ estimation problem (2.3) can be converted into finding the estimate sequence $\{\{\check{z}(k \mid k)\}_{k=0}^N; \{\check{z}_f(k \mid k)\}_{k=0}^N; \{\check{z}_h(k_d \mid k)\}_{k=d}^N\}$ such that J_N has a minimum with respect to $\{x_0, \overline{w}\}$ and the minimum of J_N is positive. As mentioned in [25, 26], the formulated H_∞ estimation problem can be solved by employing the Krein space approach.

Introduce the following Krein space stochastic system

$$\begin{aligned} \mathbf{x}(k+1) &= A\mathbf{x}(k) + A_d\mathbf{x}(k_d) + f\left(k, \check{\mathbf{z}}_f(k \mid k), \mathbf{u}(k)\right) \\ &\quad + h(k, \check{\mathbf{z}}_h(k_d \mid k), \mathbf{u}(k)) + \overline{B}\overline{\mathbf{w}}(k), \\ \mathbf{y}(k) &= C\mathbf{x}(k) + \mathbf{v}(k), \\ \check{\mathbf{z}}_f(k \mid k) &= F\mathbf{x}(k) + \mathbf{v}_{z_f}(k), \\ \check{\mathbf{z}}(k \mid k) &= L\mathbf{x}(k) + \mathbf{v}_z(k), \\ \check{\mathbf{z}}_h(k_d \mid k) &= H\mathbf{x}(k_d) + \mathbf{v}_{z_h}(k_d), \quad k \ge d, \end{aligned} \tag{3.7}$$

where $\overline{B} = [B \;\; I \;\; I]$; the initial state $\mathbf{x}_0(s)$ $(s = -d, -d+1, \ldots, 0)$ and $\overline{\mathbf{w}}(k)$, $\mathbf{v}(k)$, $\mathbf{v}_{z_f}(k)$, $\mathbf{v}_z(k)$ and $\mathbf{v}_{z_h}(k)$ are mutually uncorrelated white noises with zero means and known covariance matrices $\overline{\Pi}(s)$, $Q_{\overline{w}}(k) = I$, $Q_v(k) = I$, $Q_{v_{z_f}}(k) = -\alpha^{-2}I$, $Q_{v_z}(k) = -\gamma^2 I$, and $Q_{v_{zh}}(k) = -\beta^{-2}I$; $\check{\mathbf{z}}_f(k \mid k)$, $\check{\mathbf{z}}(k \mid k)$ and $\check{\mathbf{z}}_h(k_d \mid k)$ are regarded as the imaginary measurement at time k for the linear combination $F\mathbf{x}(k)$, $L\mathbf{x}(k)$, and $H\mathbf{x}(k_d)$, respectively.

to Contents, page 1

Let

$$\mathbf{y}_z(k) = \begin{cases} \left[\mathbf{y}^T(k) \;\; \check{\mathbf{z}}_m^T(k \mid k)\right]^T, & 0 \le k < d, \\ \left[\mathbf{y}^T(k) \;\; \check{\mathbf{z}}_m^T(k \mid k) \;\; \check{\mathbf{z}}_h^T(k_d \mid k)\right]^T, & k \ge d, \end{cases}$$

$$\mathbf{v}_{z,a}(k) = \begin{cases} \left[\mathbf{v}^T(k) \;\; \mathbf{v}_{z_f}^T(k) \;\; \mathbf{v}_z^T(k)\right]^T, & 0 \le k < d, \\ \left[\mathbf{v}^T(k) \;\; \mathbf{v}_{z_f}^T(k) \;\; \mathbf{v}_z^T(k) \;\; \mathbf{v}_{z_h}^T(k_d)\right]^T, & k \ge d, \end{cases} \tag{3.8}$$

$$\check{\mathbf{z}}_m(k \mid k) = \left[\check{\mathbf{z}}_f^T(k \mid k) \;\; \check{\mathbf{z}}^T(k \mid k)\right]^T.$$

Definition 3.1. The estimator $\hat{\mathbf{y}}(i \mid i-1)$ denotes the optimal estimation of $\mathbf{y}(i)$ given the observation $\mathcal{L}\{\{\mathbf{y}_z(j)\}_{j=0}^{i-1}\}$; the estimator $\hat{\mathbf{z}}_m(i \mid i)$ denotes the optimal estimation of $\check{\mathbf{z}}_m(i \mid i)$ given the observation $\mathcal{L}\{\{\mathbf{y}_z(j)\}_{j=0}^{i-1}; \mathbf{y}(i)\}$; the estimator $\hat{\mathbf{z}}_h(i_d \mid i)$ denotes the optimal estimation of $\check{\mathbf{z}}_h(i_d \mid i)$ given the observation $\mathcal{L}\{\{\mathbf{y}_z(j)\}_{j=0}^{i-1}; \mathbf{y}(i), \check{\mathbf{z}}_m(i \mid i)\}$.

Furthermore, introduce the following stochastic vectors and the corresponding covariance matrices

$$\begin{aligned} \tilde{\mathbf{y}}(i \mid i-1) &= \mathbf{y}(i) - \hat{\mathbf{y}}(ii-1), & R_{\tilde{y}}(ii-1) &= \langle \tilde{\mathbf{y}}(ii-1), \tilde{\mathbf{y}}(ii-1)\rangle, \\ \tilde{\mathbf{z}}_m(i \mid i) &= \check{\mathbf{z}}_m(ii) - \hat{\mathbf{z}}_m(ii), & R_{\tilde{z}_m}(ii) &= \langle \tilde{\mathbf{z}}_m(ii), \tilde{\mathbf{z}}_m(ii)\rangle, \\ \tilde{\mathbf{z}}_h(i_d \mid i) &= \check{\mathbf{z}}_h(i_d i) - \hat{\mathbf{z}}_h(i_d i), & R_{\tilde{z}_h}(i_d i) &= \langle \tilde{\mathbf{z}}_h(i_d i), \tilde{\mathbf{z}}_h(i_d i)\rangle. \end{aligned} \tag{3.9}$$

And, denote

$$\tilde{\mathbf{y}}_z(i) = \begin{cases} \left[\tilde{\mathbf{y}}^T(i \mid i-1) \;\; \tilde{\mathbf{z}}_m^T(i \mid i)\right]^T, & 0 \le i < d, \\ \left[\tilde{\mathbf{y}}^T(i \mid i-1) \;\; \tilde{\mathbf{z}}_m^T(i \mid i) \;\; \tilde{\mathbf{z}}_h^T(i_d \mid i)\right]^T, & i \ge d, \end{cases} \tag{3.10}$$

$$R_{\tilde{y}_z}(i) = \langle \tilde{\mathbf{y}}_z(i), \tilde{\mathbf{y}}_z(i)\rangle.$$

For calculating the minimum of J_N, we present the following Lemma 3.2.

Lemma 3.2. *$\{\{\tilde{\mathbf{y}}_z(i)\}_{i=0}^{k}\}$ is the innovation sequence which spans the same linear space as that of $\mathcal{L}\{\{\mathbf{y}_z(i)\}_{i=0}^{k}\}$.*

Proof. From Definition 3.1 and (3.9), $\tilde{\mathbf{y}}(i \mid i-1)$, $\tilde{\mathbf{z}}_m(i \mid i)$ and $\tilde{\mathbf{z}}_h(i_d \mid i)$ are the linear combination of the observation sequence $\{\{\mathbf{y}_z(j)\}_{j=0}^{i-1}; \mathbf{y}(i)\}$, $\{\{\mathbf{y}_z(j)\}_{j=0}^{i-1};\ \mathbf{y}(i), \check{\mathbf{z}}_m(i \mid i)\}$, and $\{\{\mathbf{y}_z(j)\}_{j=0}^{i}\}$, respectively. Conversely, $\mathbf{y}(i)$, $\check{\mathbf{z}}_m(i \mid i)$ and $\check{\mathbf{z}}_h(i_d \mid i)$ can be given by the linear combination of $\{\{\tilde{\mathbf{y}}_z(j)\}_{j=0}^{i-1}; \tilde{\mathbf{y}}(i \mid i-1)\}$, $\{\{\tilde{\mathbf{y}}_z(j)\}_{j=0}^{i-1}; \tilde{\mathbf{y}}(i \mid i-1), \tilde{\mathbf{z}}_m(i \mid i)\}$ and $\{\{\tilde{\mathbf{y}}_z(j)\}_{j=0}^{i}\}$, respectively. Hence,

$$\mathcal{L}\left\{\{\tilde{\mathbf{y}}_z(i)\}_{i=0}^{k}\right\} = \mathcal{L}\left\{\{\mathbf{y}_z(i)\}_{i=0}^{k}\right\}. \tag{3.11}$$

to Contents, page 1

It is also shown by (3.9) that $\tilde{\mathbf{y}}(i \mid i-1)$, $\tilde{\mathbf{z}}_m(i \mid i)$ and $\tilde{\mathbf{z}}_h(i_d \mid i)$ satisfy

$$\begin{aligned} \tilde{\mathbf{y}}(i \mid i-1) &\perp \mathcal{L}\Big\{ \{\mathbf{y}_z(j)\}_{j=0}^{i-1} \Big\}, \\ \tilde{\mathbf{z}}_m(i \mid i) &\perp \mathcal{L}\Big\{ \{\mathbf{y}_z(j)\}_{j=0}^{i-1}; \mathbf{y}(i) \Big\}, \\ \tilde{\mathbf{z}}_h(i_d \mid i) &\perp \mathcal{L}\Big\{ \{\mathbf{y}_z(j)\}_{j=0}^{i-1}; \mathbf{y}(i), \tilde{\mathbf{z}}_m(i \mid i) \Big\}. \end{aligned} \tag{3.12}$$

Consequently,

$$\begin{aligned} \tilde{\mathbf{y}}(i \mid i-1) &\perp \mathcal{L}\Big\{ \{\tilde{\mathbf{y}}_z(j)\}_{j=0}^{i-1} \Big\}, \\ \tilde{\mathbf{z}}_m(i \mid i) &\perp \mathcal{L}\Big\{ \{\tilde{\mathbf{y}}_z(j)\}_{j=0}^{i-1}; \tilde{\mathbf{y}}(i \mid i-1) \Big\}, \\ \tilde{\mathbf{z}}_h(i_d \mid i) &\perp \mathcal{L}\Big\{ \{\tilde{\mathbf{y}}_z(j)\}_{j=0}^{i-1}; \tilde{\mathbf{y}}(i \mid i-1), \tilde{\mathbf{z}}_m(i \mid i) \Big\}. \end{aligned} \tag{3.13}$$

This completes the proof. □

Now, an existence condition and a solution to the minimum of J_N are derived as follows.

Theorem 3.3. *Consider system* (2.1), *given a scalar* $\gamma > 0$ *and the positive definite matrix* $\Pi(k)$ $(k = -d, -d+1, \ldots, 0)$, *then* J_N *has the minimum if only if*

$$\begin{aligned} R_{\tilde{y}}(k \mid k-1) &> 0, \quad 0 \le k \le N, \\ R_{\tilde{z}_m}(k \mid k) &< 0, \quad 0 \le k \le N, \\ R_{\tilde{z}_h}(k_d \mid k) &< 0, \quad d \le k \le N. \end{aligned} \tag{3.14}$$

In this case the minimum value of J_N *is given by*

$$\begin{aligned} \min J_N = {} & \sum_{k=0}^{N} \tilde{y}^T(k \mid k-1) R_{\tilde{y}}^{-1}(k \mid k-1) \tilde{y}(k \mid k-1) + \sum_{k=0}^{N} \tilde{z}_m^T(k \mid k) R_{\tilde{z}_m}^{-1}(k \mid k) \tilde{z}_m(k \mid k) \\ & + \sum_{k=d}^{N} \tilde{z}_h^T(k_d \mid k) R_{\tilde{z}_h}^{-1}(k_d \mid k) \tilde{z}_h(k_d \mid k), \end{aligned} \tag{3.15}$$

where

$$\begin{aligned} \tilde{y}(k \mid k-1) &= y(k) - \hat{y}(k \mid k-1), \\ \tilde{z}_m(k \mid k) &= \check{z}_m(k \mid k) - \hat{z}_m(k \mid k), \\ \tilde{z}_h(k_d \mid k) &= \check{z}_h(k_d \mid k) - \hat{z}_h(k_d \mid k), \\ \check{z}_m(k \mid k) &= \begin{bmatrix} \check{z}_f^T(k \mid k) & \check{z}^T(k \mid k) \end{bmatrix}^T, \end{aligned} \tag{3.16}$$

to Contents, page 1

$\hat{y}(k \mid k-1)$ is obtained from the Krein space projection of $\mathbf{y}(k)$ onto $\mathcal{L}\{\{\mathbf{y}_z(j)\}_{j=0}^{k-1}\}$, $\hat{z}_m(k \mid k)$ is obtained from the Krein space projection of $\check{\mathbf{z}}_m(k \mid k)$ onto $\mathcal{L}\{\{\mathbf{y}_z(j)\}_{j=0}^{k-1};\ \mathbf{y}(k)\}$, and $\hat{z}_h(k_d \mid k)$ is obtained from the Krein space projection of $\check{\mathbf{z}}_h(k_d \mid k)$ onto $\mathcal{L}\{\{\mathbf{y}_z(j)\}_{j=0}^{k-1};\ \mathbf{y}(k), \check{\mathbf{z}}_m(k \mid k)\}$.

Proof. Based on the definition (3.2) and (3.3), the state equation in system (2.1) can be rewritten as

$$
\begin{aligned}
x(k+1) = {} & Ax(k) + A_d x(k_d) + f\left(k, \check{z}_f(k \mid k), u(k)\right) \\
& + h(k, \check{z}_h(k_d \mid k), u(k)) + \overline{B}\overline{w}(k).
\end{aligned} \tag{3.17}
$$

In this case, it is assumed that $f(k, \check{z}_f(k \mid k), u(k))$ and $h(k, \check{z}_h(k_d \mid k), u(k))$ are known at time k. Then, we define

$$
y_z(k) = \begin{cases} \left[y^T(k) \ \ \check{z}_f^T(k \mid k) \ \ \check{z}^T(k \mid k)\right]^T, & 0 \le k < d, \\ \left[y^T(k) \ \ \check{z}_f^T(k \mid k) \ \ \check{z}^T(k \mid k) \ \ \check{z}_h^T(k_d \mid k)\right]^T, & k \ge d. \end{cases} \tag{3.18}
$$

By introducing an augmented state

$$
x_a(k) = \left[x^T(k) \ \ x^T(k-1) \ \ \cdots \ \ x^T(k-d)\right]^T, \tag{3.19}
$$

we obtain an augmented state-space model

$$
\begin{aligned}
x_a(k+1) &= A_a x_a(k) + B_{u,a}\overline{u}(k) + \overline{B}_a \overline{w}(k), \\
y_z(k) &= C_{z,a}(k) x_a(k) + v_{z,a}(k),
\end{aligned} \tag{3.20}
$$

where

$$
A_a = \begin{bmatrix} A & 0 & \cdots & 0 & A_d \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}, \qquad B_{u,a} = \begin{bmatrix} I & I \\ 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}, \qquad \overline{B}_a = \begin{bmatrix} \overline{B} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},
$$

to Contents, page 1

$$C_{z,a}(k) = \begin{cases} \begin{bmatrix} C & 0 & \cdots & 0 \\ F & 0 & \cdots & 0 \\ L & 0 & \cdots & 0 \end{bmatrix}, & 0 \le k < d, \\ \begin{bmatrix} C & 0 & \cdots & 0 \\ F & 0 & \cdots & 0 \\ L & 0 & \cdots & 0 \\ 0 & \cdots & 0 & H \end{bmatrix}, & k \ge d, \end{cases}$$

$$v_{z,a}(k) = \begin{cases} \left[v^T(k) \;\; v_{z_f}^T(k) \;\; v_z^T(k) \right]^T, & 0 \le k < d, \\ \left[v^T(k) \;\; v_{z_f}^T(k) \;\; v_z^T(k) \;\; v_{z_h}^T(k_d) \right]^T, & k \ge d, \end{cases}$$

$$\overline{u}(k) = \left[f^T(k, \check{z}_f(k \mid k), u(k)) \;\; h^T(k, \check{z}_h(k_d \mid k), u(k)) \right]^T. \tag{3.21}$$

Additionally, we can rewrite J_N as

$$J_N = \begin{bmatrix} x_a(0) \\ \overline{w}_N \\ v_{z,aN} \end{bmatrix}^T \begin{bmatrix} P_a(0) & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & Q_{v_{z,aN}} \end{bmatrix}^{-1} \begin{bmatrix} x_a(0) \\ \overline{w}_N \\ v_{z,aN} \end{bmatrix}, \tag{3.22}$$

where

$$\begin{aligned} P_a(0) &= \operatorname{diag}\left\{ \overline{\Pi}(0), \overline{\Pi}(-1), \ldots, \overline{\Pi}(-d) \right\}, \\ \overline{w}_N &= \left[\overline{w}^T(0) \;\; \overline{w}^T(1) \;\; \cdots \;\; \overline{w}^T(N) \right]^T, \\ v_{z,aN} &= \left[v_{z,a}^T(0) \;\; v_{z,a}^T(1) \;\; \cdots \;\; v_{z,a}^T(N) \right]^T, \\ Q_{v_{z,aN}} &= \operatorname{diag}\{ Q_{v_{z,a}}(0), Q_{v_{z,a}}(1), \ldots, Q_{v_{z,a}}(N) \}, \\ Q_{v_{z,a}}(k) &= \begin{cases} \operatorname{diag}\{ I, -\gamma^2, -\alpha^{-2} \}, & 0 \le k < d, \\ \operatorname{diag}\{ I, -\gamma^2, -\alpha^{-2}, -\beta^{-2} \}, & k \ge d. \end{cases} \end{aligned} \tag{3.23}$$

Define the following state transition matrix

$$\begin{aligned} \Phi(k+1, m) &= A_a \Phi(k, m), \\ \Phi(m, m) &= I, \end{aligned} \tag{3.24}$$

to Contents, page 1

and let

$$y_{zN} = \left[y_z^T(0)\ \ y_z^T(1)\ \ \cdots\ \ y_z^T(N)\right]^T,$$
$$\overline{u}_N = \left[\overline{u}^T(0)\ \ \overline{u}^T(1)\ \ \cdots\ \ \overline{u}^T(N)\right]^T. \tag{3.25}$$

Using (3.20) and (3.24), we have

$$y_{zN} = \Psi_{0N} x_a(0) + \Psi_{\overline{u}N}\overline{u}_N + \Psi_{\overline{w}N}\overline{w}_N + v_{z,aN}, \tag{3.26}$$

where

$$\Psi_{0N} = \begin{bmatrix} C_{z,a}(0)\Phi(0,0) \\ C_{z,a}(1)\Phi(1,0) \\ \vdots \\ C_{z,a}(N)\Phi(N,0) \end{bmatrix}, \quad \Psi_{\overline{u}N} = \begin{bmatrix} \varphi_{00} & \varphi_{01} & \cdots & \varphi_{0N} \\ \varphi_{10} & \varphi_{11} & \cdots & \varphi_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{N0} & \varphi_{N1} & \cdots & \varphi_{NN} \end{bmatrix},$$
$$\varphi_{ij} = \begin{cases} C_{z,a}(i)\Phi(i,j+1)B_{u,a}, & i > j, \\ 0, & i \leq j. \end{cases} \tag{3.27}$$

The matrix $\Psi_{\overline{w}N}$ is derived by replacing $B_{u,a}$ in $\Psi_{\overline{u}N}$ with $\overline{B}_a$.

Thus, J_N can be reexpressed as

$$J_N = \begin{bmatrix} x_a(0) \\ \overline{w}_N \\ \overline{y}_{zN} \end{bmatrix}^T \left\{ \Gamma_N \begin{bmatrix} P_a(0) & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & Q_{v_{z,aN}} \end{bmatrix} \Gamma_N^T \right\}^{-1} \begin{bmatrix} x_a(0) \\ \overline{w}_N \\ \overline{y}_{zN} \end{bmatrix}, \tag{3.28}$$

where

$$\overline{y}_{zN} = y_{zN} - \Psi_{\overline{u}N}\overline{u}_N,$$
$$\Gamma_N = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ \Psi_{0N} & \Psi_{\overline{w}N} & I \end{bmatrix}. \tag{3.29}$$

Considering the Krein space stochastic system defined by (3.7) and state transition matrix (3.24), we have

$$\mathbf{y}_{zN} = \Psi_{0N}\mathbf{x}_a(0) + \Psi_{\overline{u}N}\overline{\mathbf{u}}_N + \Psi_{\overline{w}N}\overline{\mathbf{w}}_N + \mathbf{v}_{z,aN}, \tag{3.30}$$

where matrices Ψ_{0N}, $\Psi_{\overline{u}N}$, and $\Psi_{\overline{w}N}$ are the same as given in (3.26), vectors $\mathbf{y}_{zN}$ and $\overline{\mathbf{u}}_N$ are, respectively, defined by replacing Euclidean space element y_z and $\overline{u}$ in y_{zN} and $\overline{u}_N$ given

to Contents, page 1

by (3.25) with the Krein space element $\mathbf{y}_z$ and $\overline{\mathbf{u}}$, vectors $\overline{\mathbf{w}}_N$ and $\mathbf{v}_{z,aN}$ are also defined by replacing Euclidean space element $\overline{w}$ and $v_{z,a}$ in $\overline{w}_N$ and $v_{z,aN}$ given by (3.23) with the Krein space element $\overline{\mathbf{w}}$ and $\mathbf{v}_{z,a}$, and vector $\mathbf{x}_a(0)$ is given by replacing Euclidean space element x in $x_a(k)$ given by (3.19) with the Krein space element $\mathbf{x}$ when $k = 0$.

Using the stochastic characteristic of $\mathbf{x}_a(0)$, $\overline{\mathbf{w}}_N$ and $\mathbf{v}_{z,a}$, we have

$$J_N = \begin{bmatrix} x_a(0) \\ \overline{w}_N \\ \overline{y}_{zN} \end{bmatrix}^T \left\langle \begin{bmatrix} \mathbf{x}_a(0) \\ \overline{\mathbf{w}}_N \\ \overline{\mathbf{y}}_{zN} \end{bmatrix}, \begin{bmatrix} \mathbf{x}_a(0) \\ \overline{\mathbf{w}}_N \\ \overline{\mathbf{y}}_{zN} \end{bmatrix} \right\rangle^{-1} \begin{bmatrix} x_a(0) \\ \overline{w}_N \\ \overline{y}_{zN} \end{bmatrix}, \tag{3.31}$$

where $\overline{\mathbf{y}}_{zN} = \mathbf{y}_{zN} - \Psi_{\overline{u}N}\overline{\mathbf{u}}_N$.

In the light of Theorem 2.4.2 and Lemma 2.4.3 in [26], J_N has a minimum over $\{x_a(0), \overline{w}_N\}$ if and only if $R_{\overline{y}_{zN}} = \langle \overline{\mathbf{y}}_{zN}, \overline{\mathbf{y}}_{zN} \rangle$ and $Q_{v_{z,aN}} = \langle \mathbf{v}_{z,aN}, \mathbf{v}_{z,aN} \rangle$ have the same inertia. Moreover, the minimum of J_N is given by

$$\min J_N = \overline{y}_{zN}^T R_{\overline{y}_{zN}}^{-1} \overline{y}_{zN}. \tag{3.32}$$

On the other hand, applying the Krein space projection formula, we have

$$\overline{\mathbf{y}}_{zN} = \Theta_N \tilde{\mathbf{y}}_{zN}, \tag{3.33}$$

where

$$\tilde{\mathbf{y}}_{zN} = \begin{bmatrix} \tilde{\mathbf{y}}_z^T(0) & \tilde{\mathbf{y}}_z^T(1) & \cdots & \tilde{\mathbf{y}}_z^T(N) \end{bmatrix}^T,$$

$$\Theta_N = \begin{bmatrix} \theta_{00} & \theta_{01} & \cdots & \theta_{0N} \\ \theta_{10} & \theta_{11} & \cdots & \theta_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{N0} & \theta_{N1} & \cdots & \theta_{NN} \end{bmatrix},$$

to Contents, page 1

$$\theta_{ij} = \begin{cases} \langle \overline{\mathbf{y}}_z(i), \widetilde{\mathbf{y}}_z(j) \rangle R_{\widetilde{y}_z}^{-1}(j), & i > j \geq 0, \\ \begin{bmatrix} I & 0 \\ m_1 & I \end{bmatrix}, & d > i = j \geq 0, \\ \begin{bmatrix} I & 0 & 0 \\ m_1 & I & 0 \\ m_2 & m_3 & I \end{bmatrix}, & i = j \geq d, \\ 0, & 0 \leq i < j, \end{cases}$$

$$m_1 = \left\langle \overline{\check{\mathbf{z}}}_m(i \mid i), \widetilde{\mathbf{y}}(j \mid j-1) \right\rangle R_{\widetilde{y}}^{-1}(j \mid j-1),$$

$$m_2 = \left\langle \overline{\check{\mathbf{z}}}_h(i_d \mid i), \widetilde{\mathbf{y}}(j \mid j-1) \right\rangle R_{\widetilde{y}}^{-1}(j \mid j-1),$$

$$m_3 = \left\langle \overline{\check{\mathbf{z}}}_h(i_d \mid i), \widetilde{\mathbf{z}}_m(j \mid j) \right\rangle R_{\widetilde{z}_m}^{-1}(j \mid j),$$

$$\overline{\mathbf{y}}_z(i) = \mathbf{y}_z(i) - \sum_{j=0}^{N} \varphi_{ij} \overline{\mathbf{u}}(j),$$

$$\overline{\check{\mathbf{z}}}_m(i \mid i) = \check{\mathbf{z}}_m(i \mid i) - \sum_{j=0}^{N} \varphi_{m,ij} \overline{\mathbf{u}}(j),$$

$$\overline{\check{\mathbf{z}}}_h(i_d \mid i) = \check{\mathbf{z}}_h(i_d \mid i) - \sum_{j=0}^{N} \varphi_{h,ij} \overline{\mathbf{u}}(j),$$

(3.34)

where $\varphi_{m,ij}$ is derived by replacing $C_{z,a}$ in φ_{ij} with $\left[\begin{smallmatrix} F & 0 & \cdots & 0 \\ L & 0 & \cdots & 0 \end{smallmatrix}\right]$, $\varphi_{h,ij}$ is derived by replacing $C_{z,a}$ in φ_{ij} with $[0 \ \ 0 \ \ \cdots \ \ H]$ Furthermore, it follows from (3.33) that

$$R_{\overline{y}_{zN}} = \Theta_N R_{\widetilde{y}_{zN}} \Theta_N^T, \qquad \overline{y}_{zN} = \Theta_N \widetilde{y}_{zN}, \tag{3.35}$$

where

$$R_{\widetilde{y}_{zN}} = \langle \widetilde{\mathbf{y}}_{zN}, \widetilde{\mathbf{y}}_{zN} \rangle,$$

$$\widetilde{y}_{zN} = \left[\widetilde{y}_z^T(0) \ \ \widetilde{y}_z^T(1) \ \ \cdots \ \ \widetilde{y}_z^T(N)\right]^T,$$

$$\widetilde{y}_z(i) = \begin{cases} \left[\widetilde{y}^T(i \mid i-1) \ \ \widetilde{z}_m^T(i \mid i)\right]^T, & 0 \leq i < d, \\ \left[\widetilde{y}^T(i \mid i-1) \ \ \widetilde{z}_m^T(i \mid i) \ \ \widetilde{z}_h^T(i_d \mid i)\right]^T, & i \geq d. \end{cases} \tag{3.36}$$

to Contents, page 1

Since matrix Θ_N is nonsingular, it follows from (3.35) that $R_{\overline{y}_{zN}}$ and $R_{\widetilde{y}_{zN}}$ are congruent, which also means that $R_{\overline{y}_{zN}}$ and $R_{\widetilde{y}_{zN}}$ have the same inertia. Note that both $R_{\widetilde{y}_{zN}}$ and $Q_{v_{z,aN}}$ are block-diagonal matrices, and

$$R_{\widetilde{y}_z}(k) = \begin{cases} \operatorname{diag}\left\{R_{\widetilde{y}}(k \mid k-1), R_{\widetilde{z}_m}(k \mid k)\right\}, & 0 \le k < d, \\ \operatorname{diag}\left\{R_{\widetilde{y}}(k \mid k-1), R_{\widetilde{z}_m}(k \mid k), R_{\widetilde{z}_h}(k_d \mid k)\right\}, & k \le d, \end{cases} \tag{3.37}$$

$Q_{v_{z,a}}(k)$ is given by (3.23). It follows that $R_{\widetilde{y}_{zN}}$ and $Q_{v_{z,aN}}$ have the same inertia if and only if $R_{\widetilde{y}}(k \mid k-1) > 0$ $(0 \le k \le N)$, $R_{\widetilde{z}_m}(k \mid k) < 0$ $(0 \le k \le N)$ and $R_{\widetilde{z}_h}(k_d \mid k) < 0$ $(d \le k \le N)$.

Therefore, J_N subject to system (2.1) with Lipschitz conditions (2.2) has the minimum if and only if $R_{\widetilde{y}}(k \mid k-1) > 0$ $(0 \le k \le N)$, $R_{\widetilde{z}_m}(k \mid k) < 0$ $(0 \le k \le N)$ and $R_{\widetilde{z}_h}(k_d \mid k) < 0$ $(d \le k \le N)$. Moreover, the minimum value of J_N can be rewritten as

$$\begin{aligned} \min J_N &= \overline{y}_{zN}^T R_{\overline{y}_{zN}}^{-1} \overline{y}_{zN} = \widetilde{y}_{zN}^T R_{\widetilde{y}_{zN}}^{-1} \widetilde{y}_{zN} \\ &= \sum_{k=0}^{N} \widetilde{y}^T(k \mid k-1) R_{\widetilde{y}}^{-1}(k \mid k-1)\widetilde{y}(k \mid k-1) + \sum_{k=0}^{N} \widetilde{z}_m^T(k \mid k) R_{\widetilde{z}_m}^{-1}(k \mid k)\widetilde{z}_m(k \mid k) \\ &\quad + \sum_{k=d}^{N} \widetilde{z}_h^T(k_d \mid k) R_{\widetilde{z}_h}^{-1}(k_d \mid k)\widetilde{z}_h(k_d \mid k). \end{aligned} \tag{3.38}$$

The proof is completed. □

Remark 3.4. Due to the built innovation sequence $\{\{\widetilde{\mathbf{y}}_z(i)\}_{i=0}^{k}\}$ in Lemma 3.2, the form of the minimum on indefinite quadratic form J_N is different from the one given in [26–28]. It is shown from (3.15) that the estimation errors $\widetilde{y}(k \mid k-1)$, $\widetilde{z}_m(k \mid k)$ and $\widetilde{z}_h(k_d \mid k)$ are mutually uncorrelated, which will make the design of H_∞ estimator much easier than the one given in [26–28].

3.2. *Solution of the H_∞ Estimation Problem*

In this subsection, the Kalman-like recursive H_∞ estimator is presented by using orthogonal projection in the Krein space.

Denote

$$\begin{aligned} \mathbf{y}_0(i) &= \mathbf{y}(i), \\ \mathbf{y}_1(i) &= \left[\mathbf{y}^T(i) \;\; \check{\mathbf{z}}_m^T(i \mid i)\right]^T, \\ \mathbf{y}_2(i) &= \left[\mathbf{y}^T(i) \;\; \check{\mathbf{z}}_m^T(i \mid i) \;\; \check{\mathbf{z}}_h^T(i \mid i+d)\right]^T. \end{aligned} \tag{3.39}$$

to Contents, page 1

Observe from (3.8), we have

$$\mathcal{L}\left\{\{\mathbf{y}_z(i)\}_{i=0}^{j}\right\} = \mathcal{L}\left\{\{\mathbf{y}_1(i)\}_{i=0}^{j}\right\}, \quad 0 \le j < d,$$
$$\mathcal{L}\left\{\{\mathbf{y}_z(i)\}_{i=0}^{j}\right\} = \mathcal{L}\left\{\{\mathbf{y}_2(i)\}_{i=0}^{j_d}\right\}; \left\{\{\mathbf{y}_1(i)\}_{i=j_d+1}^{j}\right\}, \quad j \ge d. \tag{3.40}$$

Definition 3.5. Given $k \ge d$, the estimator $\hat{\xi}(i \mid j, 2)$ for $0 \le j < k_d$ denotes the optimal estimate of $\xi(i)$ given the observation $\mathcal{L}\{\{\mathbf{y}_2(s)\}_{s=0}^{j}\}$, and the estimator $\hat{\xi}(i \mid j, 1)$ for $k_d \le j \le k$ denotes the optimal estimate of $\xi(i)$ given the observation $\mathcal{L}\{\{\mathbf{y}_2(s)\}_{s=0}^{k_d-1}; \{\mathbf{y}_1(\tau)\}_{\tau=k_d}^{j}\}$. For simplicity, we use $\hat{\xi}(i, 2)$ to denote $\hat{\xi}(i \mid i-1, 2)$, and use $\hat{\xi}(i, 1)$ to denote $\hat{\xi}(i \mid i-1, 1)$ throughout the paper.

Based on the above definition, we introduce the following stochastic sequence and the corresponding covariance matrices

$$\begin{aligned}
\tilde{\mathbf{y}}_2(i,2) &= \mathbf{y}_2(i) - \hat{\mathbf{y}}_2(i,2), & R_{\tilde{y}_2}(i,2) &= \langle \tilde{\mathbf{y}}_2(i,2), \tilde{\mathbf{y}}_2(i,2) \rangle, \\
\tilde{\mathbf{y}}_1(i,1) &= \mathbf{y}_1(i) - \hat{\mathbf{y}}_1(i,1), & R_{\tilde{y}_1}(i,1) &= \langle \tilde{\mathbf{y}}_1(i,1), \tilde{\mathbf{y}}_1(i,1) \rangle, \\
\tilde{\mathbf{y}}_0(i,0) &= \mathbf{y}_0(i) - \hat{\mathbf{y}}_0(i,1), & R_{\tilde{y}_0}(i,0) &= \langle \tilde{\mathbf{y}}_0(i,0), \tilde{\mathbf{y}}_0(i,0) \rangle.
\end{aligned} \tag{3.41}$$

Similar to the proof of Lemma 2.2.1 in [27], we can obtain that $\{\tilde{\mathbf{y}}_2(0,2), \ldots, \tilde{\mathbf{y}}_2(k_d-1,2); \tilde{\mathbf{y}}_1(k_d,1), \ldots, \tilde{\mathbf{y}}_1(k-1,1)\}$ is the innovation sequence which is a mutually uncorrelated white noise sequence and spans the same linear space as $\mathcal{L}\{\mathbf{y}_2(0), \ldots, \mathbf{y}_2(k_d-1); \mathbf{y}_1(k_l), \ldots, \mathbf{y}_1(k-1)\}$ or equivalently $\mathcal{L}\{\mathbf{y}_z(0), \ldots, \mathbf{y}_z(k-1)\}$.

Applying projection formula in the Krein space, $\hat{\mathbf{x}}(i,2)$ $(i = 0, 1, \ldots, k_d)$ is computed recursively as

$$\tag{3.42}$$

$$\begin{aligned}
\hat{\mathbf{x}}(i+1,2) &= \sum_{j=0}^{i} \langle \mathbf{x}(i+1), \tilde{\mathbf{y}}_2(j,2) \rangle R_{\tilde{y}_2}^{-1}(j,2) \tilde{\mathbf{y}}_2(j,2) \\
&= A\hat{\mathbf{x}}(i \mid i,2) + A_d \hat{\mathbf{x}}(i_d \mid i,2) + f\left(i, \check{\mathbf{z}}_f(i \mid i), \mathbf{u}(i)\right) \\
&\quad + h(i, \check{\mathbf{z}}_h(i_d \mid i), \mathbf{u}(i)), \quad i = 0, 1, \ldots, k_d - 1, \\
\hat{\mathbf{x}}(\tau,2) &= 0, \quad (\tau = -d, -d+1, \ldots, 0).
\end{aligned} \tag{3.43}$$

Note that

$$\begin{aligned}
\hat{\mathbf{x}}(i \mid i,2) &= \hat{\mathbf{x}}(i,2) + P_2(i,i) C_2^T R_{\tilde{y}_2}^{-1}(i,2) \tilde{\mathbf{y}}_2(i,2), \\
\hat{\mathbf{x}}(i_d \mid i,2) &= \hat{\mathbf{x}}(i_d,2) + \sum_{j=i_d}^{i} P_2(i_d,j) C_2^T R_{\tilde{y}_2}^{-1}(j,2) \tilde{\mathbf{y}}_2(j,2),
\end{aligned} \tag{3.44}$$

to Contents, page 1

where

$$\begin{aligned}
C_2 &= \left[C^T \quad F^T \quad L^T \quad H^T\right]^T, \\
P_2(i,j) &= \langle \mathbf{e}(i,2), \mathbf{e}(j,2) \rangle, \\
\mathbf{e}(i,2) &= \mathbf{x}(i) - \widehat{\mathbf{x}}(i,2), \\
R_{\widetilde{y}_2}(i,2) &= C_2 P_2(i,i) C_2^T + Q_{v_2}(i), \\
Q_{v_2}(i) &= \operatorname{diag}\left\{I, -\alpha^{-2} I, -\gamma^2 I, -\beta^{-2} I\right\}.
\end{aligned} \tag{3.45}$$

Substituting (3.44) into (3.43), we have

$$\begin{aligned}
\widehat{\mathbf{x}}(i+1,2) &= A\widehat{\mathbf{x}}(i,2) + A_d \widehat{\mathbf{x}}(i_d,2) + f\left(i, \breve{\mathbf{z}}_f(i \mid i), \mathbf{u}(i)\right) + h(i, \breve{\mathbf{z}}_h(i_d \mid i), \mathbf{u}(i)) \\
&\quad + A_d \sum_{j=i_d}^{i-1} P_2(i_d,j) C_2^T R_{\widetilde{y}_2}^{-1}(j,2) \widetilde{\mathbf{y}}_2(j,2) + K_2(i) \widetilde{\mathbf{y}}_2(i,2), \\
K_2(i) &= A_d P_2(i_d,i) C_2^T R_{\widetilde{y}_2}^{-1}(i,2) + A P_2(i,i) C_2^T R_{\widetilde{y}_2}^{-1}(i,2).
\end{aligned} \tag{3.46}$$

Moreover, taking into account (3.7) and (3.46), we obtain

$$\begin{aligned}
\mathbf{e}(i+1,2) &= A\mathbf{e}(i,2) + A_d \mathbf{e}(i_d,2) + \overline{B}\overline{\mathbf{w}}(i) - K_2(i)\widetilde{\mathbf{y}}_2(i,2) \\
&\quad - A_d \sum_{j=i_d}^{i-1} P_2(i_d,j) C_2^T R_{\widetilde{y}_2}^{-1}(j,2) \widetilde{\mathbf{y}}_2(j,2), \quad i = 0,1,\ldots,k_d - 1.
\end{aligned} \tag{3.47}$$

Consequently,

$$\begin{aligned}
P_2(i-j,i+1) &= \langle \mathbf{e}(i-j,2), \mathbf{e}(i+1,2) \rangle \\
&= P_2(i-j,i) A^T + P_2^T(i_d,i-j) A_d^T - P_2(i-j,i) C_2^T K_2^T(i) \\
&\quad - \sum_{t=i-j}^{i-1} P_2(i-j,t) C_2^T R_{\widetilde{y}_2}^{-1}(t,2) C_2 P_2^T(i_d,t) A_d^T, \qquad j = 0,1,\ldots,d, \\
P_2(i+1,i+1) &= \langle \mathbf{e}(i+1,2), \mathbf{e}(i+1,2) \rangle \\
&= A P_2(i,i+1) + A_d P_2(i_d,i+1) + \overline{B} Q_{\overline{w}}(i) \overline{B}^T,
\end{aligned} \tag{3.48}$$

to Contents, page 1

where $Q_{\overline{w}}(i) = I$. Thus, $P_2(i,i)$ $(i = 0, 1, \ldots, k_d)$ can be computed recursively as

$$\begin{aligned} P_2(i-j,i+1) &= P_2(i-j,i)A^T + P_2^T(i_d,i-j)A_d^T - P_2(i-j,i)C_2^T K_2^T(i) \\ &\quad - \sum_{t=i-j}^{i-1} P_2(i-j,t)C_2^T R_{\tilde{y}_2}^{-1}(t,2)C_2 P_2^T(i_d,t)A_d^T, \\ P_2(i+1,i+1) &= AP_2(i,i+1) + A_d P_2(i_d,i+1) + \overline{B}Q_{\overline{w}}(i)\overline{B}^T, \quad j = 0,1,\ldots,d. \end{aligned} \tag{3.49}$$

Similarly, employing the projection formula in the Krein space, the optimal estimator $\widehat{\mathbf{x}}(i,1)$ $(i = k_d + 1, \ldots, k)$ can be computed by

$$\begin{aligned} \widehat{\mathbf{x}}(i+1,1) &= A\widehat{\mathbf{x}}(i,1) + A_d\widehat{\mathbf{x}}(i_d,2) + f\left(i,\check{\mathbf{z}}_f(i \mid i),\mathbf{u}(i)\right) + h(i,\check{\mathbf{z}}_h(i_d \mid i),\mathbf{u}(i)) \\ &\quad + K_1(i)\widetilde{\mathbf{y}}_1(i,1) + A_d\sum_{j=i_d}^{k_d-1} P_2(i_d,j)C_2^T R_{\tilde{y}_2}^{-1}(j,2)\widetilde{\mathbf{y}}_2(j,2) \\ &\quad + A_d\sum_{j=k_d}^{i-1} P_1(i_d,j)C_1^T R_{\tilde{y}_1}^{-1}(j,1)\widetilde{\mathbf{y}}_1(j,1), \\ \widehat{\mathbf{x}}(k_d,1) &= \widehat{\mathbf{x}}(k_d,2), \end{aligned} \tag{3.50}$$

where

$$\begin{gathered} C_1 = \left[C^T \;\; F^T \;\; L^T\right]^T, \\ P_1(i,j) = \begin{cases} \left\langle \mathbf{e}(i,2), \mathbf{e}(j,1)\right\rangle, & i < k_d, \\ \left\langle \mathbf{e}(i,1), \mathbf{e}(j,1)\right\rangle, & i \geq k_d, \end{cases} \\ \mathbf{e}(i,1) = \mathbf{x}(i) - \widehat{\mathbf{x}}(i,1), \\ R_{\tilde{y}_1}(i,1) = C_1P_1(i,i)C_1^T + Q_{v_1}(i), \\ Q_{v_1}(i) = \operatorname{diag}\left\{I, -\alpha^{-2}I, -\gamma^2 I\right\}, \\ K_1(i) = AP_1(i,i)C_1^T R_{\tilde{y}_1}^{-1}(i,1) + A_dP_1(i_d,i)C_1^T R_{\tilde{y}_1}^{-1}(i,1). \end{gathered} \tag{3.51}$$

Then, from (3.7) and (3.50), we can yield

$$\begin{aligned} \mathbf{e}(i+1,1) &= A\mathbf{e}(i,1) + A_d\mathbf{e}(i_d,2) + \overline{B}\overline{\mathbf{w}}(i) - K_1(i)\widetilde{\mathbf{y}}_1(i,1) \\ &\quad - A_d\sum_{j=i_d}^{k_d-1} P_2(i_d,j)C_2^T R_{\tilde{y}_2}^{-1}(j,2)\widetilde{\mathbf{y}}_2(j,2) \\ &\quad - A_d\sum_{j=k_d}^{i-1} P_1(i_d,j)C_1^T R_{\tilde{y}_1}^{-1}(j,1)\widetilde{\mathbf{y}}_1(j,1). \end{aligned} \tag{3.52}$$

to Contents, page 1

Thus, we obtain that

(1) if $i-j \geq k_d$, we have

$$\begin{aligned} P_1(i-j,i+1) &= \langle \mathbf{e}(i-j,1), \mathbf{e}(i+1,1) \rangle \\ &= P_1(i-j,i)A^T + P_1^T(i_d,i-j)A_d^T - P_1(i-j,i)C_1^T K_1^T(i) \\ &\quad - \sum_{t=i-j}^{i-1} P_1(i-j,t)C_1^T R_{\tilde{y}_1}^{-1}(t,1)C_1 P_1^T(i_d,t)A_d^T, \end{aligned} \tag{3.53}$$

(2) if $i-j < k_d$, we have

$$\begin{aligned} P_1(i-j,i+1) &= \langle \mathbf{e}(i-j,2), \mathbf{e}(i+1,1) \rangle \\ &= P_1(i-j,i)A^T + P_2^T(i_d,i-j)A_d^T - P_1(i-j,i)C_1^T K_1^T(i) \\ &\quad - \sum_{t=i-j}^{k_d-1} P_2(i-j,t)C_2^T R_{\tilde{y}_2}^{-1}(t,2)C_2 P_2^T(i_d,t)A_d^T \\ &\quad - \sum_{t=k_d}^{i-1} P_1(i-j,t)C_1^T R_{\tilde{y}_1}^{-1}(t,1)C_1 P_1^T(i_d,t)A_d^T, \end{aligned} \tag{3.54}$$

$$\begin{aligned} P_1(i+1,i+1) &= \langle \mathbf{e}(i-j,2), \mathbf{e}(i+1,1) \rangle \\ &= AP_1(i,i+1) + A_d P_1(i_d,i+1) + \overline{B}Q_{\overline{w}}(i)\overline{B}^T. \end{aligned} \tag{3.55}$$

It follows from (3.53), (3.54), and (3.55) that $P_1(i,i)$ $(i = k_d+1,\ldots,k)$ can be computed by

$$\begin{aligned} P_1(i-j,i+1) &= P_1(i-j,i)A^T + P_2^T(i_d,i-j)A_d^T - P_1(i-j,i)C_1^T K_1^T(i) \\ &\quad - \sum_{t=i-j}^{k_d-1} P_2(i-j,t)C_2^T R_{\tilde{y}_2}^{-1}(t,2)C_2 P_2^T(i_d,t)A_d^T \\ &\quad - \sum_{t=k_d}^{i-1} P_1(i-j,t)C_1^T R_{\tilde{y}_1}^{-1}(t,1)C_1 P_1^T(i_d,t)A_d^T, \quad i-j<k_d, \\ P_1(i-j,i+1) &= P_1(i-j,i)A^T + P_1^T(i_d,i-j)A_d^T - P_1(i-j,i)C_1^T K_1^T(i) \\ &\quad - \sum_{t=i-j}^{i-1} P_1(i-j,t)C_1^T R_{\tilde{y}_1}^{-1}(t,1)C_1 P_1^T(i_d,t)A_d^T, \quad i-j \geq k_d, \\ P_1(i+1,i+1) &= AP_1(i,i+1) + A_d P_1(i_d,i+1) + \overline{B}Q_{\overline{w}}(i)\overline{B}^T, \quad j=0,1,\ldots,d. \end{aligned} \tag{3.56}$$

to Contents, page 1

Next, according to the above analysis, $\hat{\mathbf{z}}_m(k \mid k)$ as the Krein space projections of $\check{\mathbf{z}}_m(k \mid k)$ onto $\mathcal{L}\{\{\mathbf{y}_z(j)\}_{j=0}^{k-1}; \mathbf{y}_0(k)\}$ can be computed by the following formula

$$\hat{\mathbf{z}}_m(k \mid k) = C_m\hat{\mathbf{x}}(k,1) + C_m P_1(k,k)C^T R_{\tilde{y}_0}^{-1}(k,0)\tilde{\mathbf{y}}_0(k,0), \tag{3.57}$$

where

$$\begin{gathered} C_m = \left[F^T \quad L^T\right]^T, \\ R_{\tilde{y}_0}(k,0) = CP_1(k,k)C^T + Q_v(k). \end{gathered} \tag{3.58}$$

And, $\hat{\mathbf{z}}_h(k_d \mid k)$ as the Krein space projections of $\check{\mathbf{z}}_h(k_d \mid k)$ onto $\mathcal{L}\{\{\mathbf{y}_z(j)\}_{j=0}^{k-1}; \mathbf{y}_1(k)\}$ can be computed by the following formula

$$\hat{\mathbf{z}}(k_d \mid k) = H\hat{\mathbf{x}}(k_d,1) + \sum_{j=k_d}^{k} HP_1(k_d,j)C_1^T R_{\tilde{y}_1}^{-1}(j,1)\tilde{\mathbf{y}}_1(j,1). \tag{3.59}$$

Based on Theorem 3.3 and the above discussion, we propose the following results.

Theorem 3.6. *Consider system* (2.1) *with Lipschitz conditions* (2.2), *given a scalar* $\gamma > 0$ *and matrix* $\Pi(k)$ $(k = -d, \ldots, 0)$, *then the* H_∞ *estimator that achieves* (2.3) *if*

$$\begin{gathered} R_{\tilde{y}}(k \mid k-1) > 0, \quad 0 \le k \le N, \\ R_{\tilde{z}_m}(k \mid k) < 0, \quad 0 \le k \le N, \\ R_{\tilde{z}_h}(k_d \mid k) < 0, \quad d \le k \le N, \end{gathered} \tag{3.60}$$

where

$$\begin{gathered} R_{\tilde{y}}(k \mid k-1) = R_{\tilde{y}_0}(k,0), \\ R_{\tilde{z}_m}(k \mid k) = C_m P_1(k,k)C_m^T - C_m P_1(k,k)C^T R_{\tilde{y}_0}^{-1}(k,0)CP_1(k,k)C_m^T + Q_{v_m}(k), \\ R_{\tilde{z}_h}(k_d \mid k) = HP_1(k_d,k_d)H^T - \sum_{j=k_d}^{k} HP_1(k_d,j)C_1^T R_{\tilde{y}_1}^{-1}(j,1)C_1 P_1^T(k_d,j)H^T - \beta^{-2}I, \\ Q_{v_m}(k) = \operatorname{diag}\left\{-\alpha^{-2}I, -\gamma^2 I\right\}, \end{gathered} \tag{3.61}$$

$R_{\tilde{y}_0}(k,0)$, $P_1(i,j)$, *and* $R_{\tilde{y}_1}(j,1)$ *are calculated by* (3.58), (3.56), *and* (3.51), *respectively.*

Moreover, one possible level-γ H_∞ *estimator is given by*

$$\check{z}(k \mid k) = E\hat{z}_m(k \mid k), \tag{3.62}$$

where $E = [0 \quad I]$, *and* $\hat{z}_m(k \mid k)$ *is computed by* (3.57).

to Contents, page 1

Proof. In view of Definitions 3.1 and 3.5, it follows from (3.9) and (3.41) that $R_{\tilde{y}}(k \mid k-1) = R_{\tilde{y}_0}(k,0)$. In addition, according to (3.7), (3.9), and (3.57), the covariance matrix $R_{\tilde{z}_m}(k \mid k)$ can be given by the second equality in (3.61). Similarly, based on (3.7), (3.9), and (3.59), the covariance matrix $R_{\tilde{z}_h}(k_d \mid k)$ can be obtained by the third equality in (3.61). Thus, from Theorem 3.3, it follows that J_N has a minimum if (3.60) holds.

On the other hand, note that the minimum value of J_N is given by (3.15) in Theorem 3.3 and any choice of estimator satisfying min $J_N > 0$ is an acceptable one. Therefore, Taking into account (3.60), one possible estimator can be obtained by setting $\check{z}_m(k \mid k) = \widehat{z}_m(k \mid k)$ and $\check{z}_h(k_d \mid k) = \widehat{z}_h(k_d \mid k)$. This completes the proof. □

Remark 3.7. It is shown from (3.57) and (3.59) that $\widehat{\mathbf{z}}_m(k \mid k)$ and $\widehat{\mathbf{z}}(k_d \mid k)$ are, respectively, the filtering estimate and fixed-lag smoothing of $\check{\mathbf{z}}_m(k \mid k)$ and $\check{\mathbf{z}}(k_d \mid k)$ in the Krein space. Additionally, it follows from Theorem 3.6 that $\check{z}_m(k \mid k)$ and $\check{z}_h(k_d \mid k)$ achieving the H_∞ estimation problem (2.3) can be, respectively, computed by the right side of (3.57) and (3.59). Thus, it can be concluded that the proposed results in this paper are related with both the H_2 filtering and H_2 fixed-lag smoothing in the Krein space.

Remark 3.8. Recently, the robust H_∞ observers for Lipschitz nonlinear delay-free systems with Lipschitz nonlinear additive uncertainties and time-varying parametric uncertainties have been studied in [10, 11], where the optimization of the admissible Lipschitz constant and the disturbance attenuation level are discussed simultaneously by using the multiobjective optimization technique. In addition, the sliding mode observers with H_∞ performance have been designed for Lipschitz nonlinear delay-free systems with faults (matched uncertainties) and disturbances in [8]. Although the Krein space-based robust H_∞ filter has been proposed for discrete-time uncertain linear systems in [28], it cannot be applied to solving the H_∞ estimation problem given in [10] since the considered system contains Lipschitz nonlinearity and Lipschitz nonlinear additive uncertainty. However, it is meaningful and promising in the future, by combining the algorithm given in [28] with our proposed method in this paper, to construct a Krein space-based robust H_∞ filter for discrete-time Lipschitz nonlinear systems with nonlinear additive uncertainties and time-varying parametric uncertainties.

4. A Numerical Example

Consider the system (2.1) with time delay $d = 3$ and the parameters

$$
\begin{gathered}
A = \begin{bmatrix} 0.7 & 0 \\ 0 & -0.4 \end{bmatrix}, \qquad A_d = \begin{bmatrix} -0.5 & 0 \\ 0 & 0.3 \end{bmatrix}, \qquad F = \begin{bmatrix} 0.02 & 0 \\ 0 & 0.01 \end{bmatrix}, \\
H = \begin{bmatrix} 0.03 & 0 \\ 0 & 0.02 \end{bmatrix}, \qquad B = \begin{bmatrix} 1.2 \\ 0.7 \end{bmatrix}, \qquad C = \begin{bmatrix} 1.7 & 0.9 \end{bmatrix}, \qquad L = \begin{bmatrix} 0.5 & 0.6 \end{bmatrix}, \\
f(k, Fx(k), u(k)) = \sin(Fx(k)), \qquad h(k, Hx(k_d), u(k)) = \cos(Hx(k_d)).
\end{gathered} \tag{4.1}
$$

Then we have $\alpha = \beta = 1$. Set $x(k) = [-0.2k \;\; 0.1k]^T$ $(k = -3, -2, -1, 0)$, and $\Pi(k) = I$ $(k = -3, -2, -1, 0)$. Both the system noise $w(k)$ and the measurement noise $v(k)$ are supposed to be band-limited white noise with power 0.01. By applying Theorem 3.1 in [20], we obtain the minimum disturbance attenuation level $\gamma_{\min} = 1.6164$ and the observer

to Contents, page 1

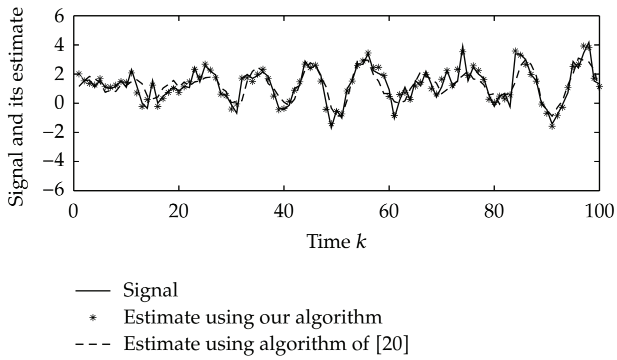


Figure 1: Signal $z(k)$ (solid), its estimate using our algorithm (star), and its estimate using algorithm in [20] (dashed).

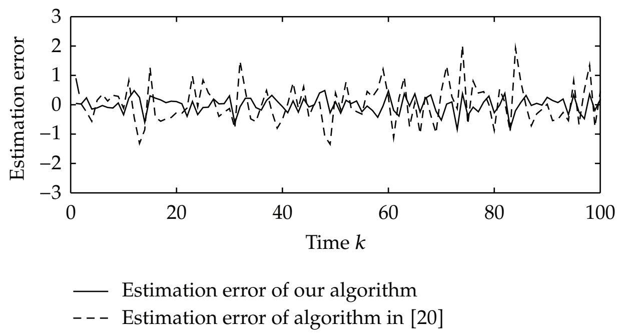


Figure 2: Estimation error of our algorithm (solid) and estimation error of algorithm in [20] (dashed).

parameter $L = [-0.3243 \;\; 0.0945]^T$ of (5) in [20]. In this numerical example, we compare our algorithm with the one given in [20] in case of $\gamma = 1.6164$. Figure 1 shows the true value of signal $z(k)$, the estimate using our algorithm, and the estimate using the algorithm given in [20]. Figure 2 shows the estimation error of our approach and the estimation error of the approach in [20]. It is shown in Figures 1 and 2 that the proposed algorithm is better than the one given in [20]. Figure 3 shows the ratios between the energy of the estimation error and input noises for the proposed H_∞ estimation algorithm. It is shown that the maximum energy ratio from the input noises to the estimation error is less than γ^2 by using our approach. Figure 4 shows the value of indefinite quadratic form J_N for the given estimation algorithm. It is shown that the value of indefinite quadratic form J_N is positive by employing the proposed algorithm in Theorem 3.6.

5. Conclusions

A recursive H_∞ filtering estimate algorithm for discrete-time Lipschitz nonlinear systems with time-delay and disturbance input is proposed. By combining the H_∞-norm estimation condition with the Lipschitz conditions on nonlinearity, the H_∞ estimation problem is converted to the positive minimum problem of indefinite quadratic form. Motivated by the observation that the minimum problem of indefinite quadratic form coincides with Kalman filtering in the Krein space, a novel Krein space-based H_∞ filtering estimate algorithm is

to Contents, page 1

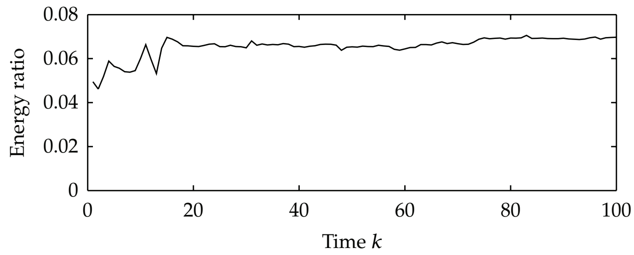


Figure 3: The energy ratio between estimation error and all input noises for the proposed H_∞ estimation algorithm.

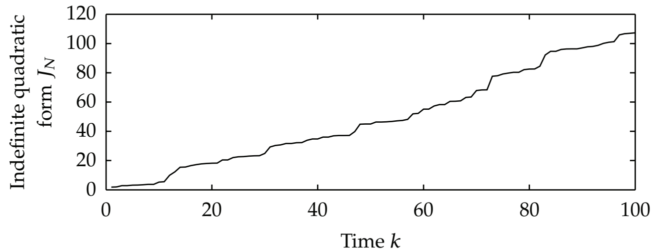


Figure 4: The value of indefinite quadratic form J_N for the given estimation algorithm.

developed. Employing projection formula and innovation analysis technology in the Krein space, the H_∞ estimator and its sufficient existence condition are presented based on Riccati-like difference equations. A numerical example is provided in order to demonstrate the performances of the proposed approach.

Future research work will extend the proposed method to investigate more general nonlinear system models with nonlinearity in observation equations. Another interesting research topic is the H_∞ multistep prediction and fixed-lag smoothing problem for time-delay Lipschitz nonlinear systems.

Acknowledgments

The authors sincerely thank the anonymous reviewers for providing valuable comments and useful suggestions aimed at improving the quality of this paper. The authors also thank the editor for the efficient and professional processing of our paper. This paper is supported by the National Natural Science Foundation of China (nos. 60774004, 61034007, 60874016) and the Independent Innovation Foundation of Shandong University, China (no. 2010JC003).

References

[1] K. Reif, F. Sonnemann, and R. Unbehauen, "An EKF-based nonlinear observer with a prescribed degree of stability," *Automatica*, vol. 34, no. 9, pp. 1119–1123, 1998.

[2] X. H. Xia and W. B. Gao, "Nonlinear observer design by observer error linearization," *SIAM Journal on Control and Optimization*, vol. 27, no. 1, pp. 199–216, 1989.

[3] C. Califano, S. Monaco, and D. Normand-Cyrot, "On the observer design in discrete-time," *Systems & Control Letters*, vol. 49, no. 4, pp. 255–265, 2003.

to Contents, page 1

[4] J. Zhang, G. Feng, and H. Xu, "Observer design for nonlinear discrete-time systems: immersion and dynamic observer error linearization techniques," *International Journal of Robust and Nonlinear Control*, vol. 20, no. 5, pp. 504–514, 2010.

[5] F. E. Thau, "Observing the state of nonlinear dynamic systems," *International Journal of Control*, vol. 17, no. 3, pp. 471–479, 1973.

[6] A. Rajamani, "Observers for Lipschitz nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 3, pp. 397–401, 1998.

[7] F. Zhu and Z. Han, "A note on observers for Lipschitz nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 10, pp. 1751–1754, 2002.

[8] R. Raoufi, H. J. Marquez, and A. S. I. Zinober, "H_∞ sliding mode observers for uncertain nonlinear Lipschitz systems with fault estimation synthesis," *International Journal of Robust and Nonlinear Control*, vol. 20, no. 16, pp. 1785–1801, 2010.

[9] W. Chen and M. Saif, "A sliding mode observer-based strategy for fault detection, isolation, and estimation in a class of Lipschitz nonlinear systems," *International Journal of Systems Science*, vol. 38, no. 12, pp. 943–955, 2007.

[10] M. Abbaszadeh and H. J. Marquez, "LMI optimization approach to robust H_∞ observer design and static output feedback stabilization for discrete-time nonlinear uncertain systems," *International Journal of Robust and Nonlinear Control*, vol. 19, no. 3, pp. 313–340, 2009.

[11] M. Abbaszadeh and H. J. Marquez, "Dynamical robust H_∞ filtering for nonlinear uncertain systems: an LMI approach," *Journal of the Franklin Institute*, vol. 347, no. 7, pp. 1227–1241, 2010.

[12] A. M. Pertew, H. J. Marquez, and Q. Zhao, "H_∞ observer design for Lipschitz nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 51, no. 7, pp. 1211–1216, 2006.

[13] A. Zemouche, M. Boutayeb, and G. I. Bara, "Observers for a class of Lipschitz systems with extension to H_∞ performance analysis," *Systems & Control Letters*, vol. 57, no. 1, pp. 18–27, 2008.

[14] A. Zemouche and M. Boutayeb, "A unified H_∞ adaptive observer synthesis method for a class of systems with both Lipschitz and monotone nonlinearities," *Systems & Control Letters*, vol. 58, no. 4, pp. 282–288, 2009.

[15] C. F. Cheng, "Design of adaptive observation schemes for Lipschitz nonlinear systems," *JSME International Journal, Series C: Mechanical Systems, Machine Elements and Manufacturing*, vol. 44, no. 3, pp. 863–866, 2001.

[16] A. Poznyak, A. Nazin, and D. Murano, "Observer matrix gain optimization for stochastic continuous time nonlinear systems," *Systems & Control Letters*, vol. 52, no. 5, pp. 377–385, 2004.

[17] M. S. Chen and C. C. Chen, "Robust nonlinear observer for Lipschitz nonlinear systems subject to disturbances," *IEEE Transactions on Automatic Control*, vol. 52, no. 12, pp. 2365–2369, 2007.

[18] D. Koenig, "Observer design for unknown input nonlinear descriptor systems via convex optimization," *IEEE Transactions on Automatic Control*, vol. 51, no. 6, pp. 1047–1052, 2006.

[19] I. R. Petersen, "Robust guaranteed cost state estimation for nonlinear stochastic uncertain systems via an IQC approach," *Systems & Control Letters*, vol. 58, no. 12, pp. 865–870, 2009.

[20] G. Lu and G. Feng, "Robust H_∞ observers for Lipschitz nonlinear discrete-time systems with time delay," *IET Control Theory & Applications*, vol. 1, no. 3, pp. 810–816, 2007.

[21] H. Gao and C. Wang, "Delay-dependent robust H_∞ and $L_2 - L_\infty$ filtering for a class of uncertain nonlinear time-delay systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 9, pp. 1661–1666, 2003.

[22] S. Xu, J. Lu, S. Zhou, and C. Yang, "Design of observers for a class of discrete-time uncertain nonlinear systems with time delay," *Journal of the Franklin Institute*, vol. 341, no. 3, pp. 295–308, 2004.

[23] A. J. Koshkouei and K. J. Burnham, "Discontinuous observers for non-linear time-delay systems," *International Journal of Systems Science*, vol. 40, no. 4, pp. 383–392, 2009.

[24] S. Ibrir, W. Xie, and C. Su, "Observer design for discrete-time systems subject to time-delay nonlinearities," *International Journal of Systems Science*, vol. 37, no. 9, pp. 629–641, 2006.

[25] H. Zhao, M. Zhong, and M. Zhang, "H_∞ fault detection for linear discrete time-varying systems with delayed state," *IET Control Theory and Applications*, vol. 4, no. 11, pp. 2303–2314, 2010.

[26] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite Quadratic Estimation and Control : A Unified Approach to H_2 and H_∞ Theorie*, Society for Industrial and Applied Mathematics, New York, NY, USA, 1999.

[27] H. Zhang and L. Xie, *Control and Estimation of Systems with Input/Output Delays*, Springer, Berlin, Germany, 2007.

[28] W. S. Ra, S. H. Jin, and J. B. Park, "Set-valued estimation approach to recursive robust H_∞ filtering," *IET Control Theory & Applications*, vol. 151, no. 6, pp. 773–782, 2004.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 154916, 23 pages
doi:10.1155/2011/154916

Research Article

Instable Trivial Solution of Autonomous Differential Systems with Quadratic Right-Hand Sides in a Cone

D. Ya. Khusainov,[1] J. Diblík,[2,3] Z. Svoboda,[2] and Z. Šmarda[2]

[1] *Department of Complex System Modeling, Faculty of Cybernetics, Taras Shevchenko National University of Kyiv, Vladimirskaya Str. 64, 01033 Kyiv, Ukraine*
[2] *Department of Mathematics, Faculty of Electrical Engineering and Communication, Technická 8, Brno University of Technology, 61600 Brno, Czech Republic*
[3] *Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Veveří 331/95, Brno University of Technology, 60200 Brno, Czech Republic*

Correspondence should be addressed to D. Ya. Khusainov, khusainov@stab.univ.kiev.ua

Received 5 October 2010; Accepted 2 November 2010

Academic Editor: Miroslava Růžičková

The present investigation deals with global instability of a general n-dimensional system of ordinary differential equations with quadratic right-hand sides. The global instability of the zero solution in a given cone is proved by Chetaev's method, assuming that the matrix of linear terms has a simple positive eigenvalue and the remaining eigenvalues have negative real parts. The sufficient conditions for global instability obtained are formulated by inequalities involving norms and eigenvalues of auxiliary matrices. In the proof, a result is used on the positivity of a general third-degree polynomial in two variables to estimate the sign of the full derivative of an appropriate function in a cone.

1. Introduction

Recently, there has been a rapidly growing interest in investigating the instability conditions of differential systems. The number of papers dealing with instability problems is rather low compared with the huge quantity of papers in which the stability of the motion of differential systems is investigated. The first results on the instability of zero solution of differential systems were obtained in a general form by Lyapunov [1] and Chetaev [2].

Further investigation on the instability of solutions of systems was carried out to weaken the conditions of the Lyapunov and Chetaev theorems for special-form systems. Some results are presented, for example, in [3–10], but instability problems are analysed only locally. For example, in [7], a linear system of ordinary differential equations in the matrix form is considered, and conditions such that the corresponding forms (of the second and the

to Contents, page 1

third power) have fixed sign in some cone of the space $\mathbb{R}^n$ are derived. To investigate this property another problem inverse to the known Lyapunov problem for the construction of Lyapunov functions is solved.

In the present paper, instability solutions of systems with quadratic right-hand sides is investigated in a cone dealing with a general n-dimensional system with quadratic right-hand sides. We assume that the matrix of linear terms has a simple positive eigenvalue and the remaining eigenvalues have negative real parts.

Unlike the previous investigations, we prove the global instability of the zero solution in a given cone and the conditions for global instability are formulated by inequalities involving norms and eigenvalues of auxiliary matrices. The main tool is the method of Chetaev and application of a suitable Chetaev-type function. A novelty in the proof of the main result (Theorem 3.1) is the utilization of a general third-order polynomial inequality of two variables to estimate the sign of the full derivative of an appropriate function along the trajectories of a given system in a cone.

In the sequel, the norms used for vectors and matrices are defined as

$$\|x\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}, \tag{1.1}$$

for a vector $x = (x_1, \ldots, x_n)^T$ and

$$\|\mathcal{F}\| = \left(\lambda_{\max}\left(\mathcal{F}^T\mathcal{F}\right)\right)^{1/2}, \tag{1.2}$$

for any $m \times n$ matrix $\mathcal{F}$. Here and throughout the paper, $\lambda_{\max}(\cdot)$ (or $\lambda_{\min}(\cdot)$) is the maximal (or minimal) eigenvalue of the corresponding symmetric and positive-semidefinite matrix $\mathcal{F}^T\mathcal{F}$ (see, e.g., [11]).

In this paper, we consider the instability of the trivial solution of a nonlinear autonomous differential system with quadratic right-hand sides

$$\dot{x}_i = \sum_{s=1}^{n} a_{is}x_s + \sum_{s,q=1}^{n} b_{sq}^i x_s x_q, \quad i = 1, \ldots, n, \tag{1.3}$$

where coefficients a_{is} and b_{sq}^i are constants. Without loss of generality, throughout this paper we assume

$$b_{sq}^i = b_{qs}^i. \tag{1.4}$$

As emphasized, for example, in [2, 10–12], system (1.3) can be written in a general vector-matrix form

$$\dot{x} = Ax + X^T Bx, \tag{1.5}$$

to Contents, page 1

where A is an $n \times n$ constant square matrix, matrix X^T is an $n \times n^2$ rectangular matrix

$$X^T = \left\{X_1^T, X_2^T, \dots, X_n^T\right\}, \tag{1.6}$$

where the entries of the $n \times n$ square matrices X_i, $i = 1, \dots, n$ are equal to zero except the ith row with entries $x^T = (x_1, x_2, \dots, x_n)$, that is,

$$X_i^T = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \\ x_1 & x_2 & \cdots & x_n \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \tag{1.7}$$

and B is a rectangular $n^2 \times n$ matrix such that

$$B^T = \{B_1, B_2, \dots, B_n\}, \tag{1.8}$$

where matrices $B_i = \{b_{sq}^i\}$, $i, s, q = 1, \dots, n$, that is, matrices

$$B_i = \begin{pmatrix} b_{11}^i & b_{12}^i & \cdots & b_{1n}^i \\ b_{21}^i & b_{22}^i & \cdots & b_{2n}^i \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1}^i & b_{n2}^i & \cdots & b_{nn}^i \end{pmatrix} \tag{1.9}$$

are $n \times n$ constant and symmetric. Representation (1.5) permits an investigation of differential systems with quadratic right-hand sides by methods of matrix analysis. Such approach was previously used, for example, in [13].

If matrix A admits one simple positive eigenvalue, the system (1.5) can be transformed, using a suitable linear transformation of the dependent variables, to the same form (1.5) but with the matrix A having the form

$$A = \begin{pmatrix} A_0 & \theta \\ \theta^T & \lambda \end{pmatrix}, \tag{1.10}$$

where A_0 is an $(n-1) \times (n-1)$ constant matrix, $\theta = (0, 0, \dots, 0)^T$ is the $(n-1)$-dimensional zero vector and $\lambda > 0$. With regard to this fact, we do not introduce new notations for the coefficients b_{sq}^i, $i, s, q = 1, 2, \dots, n$ in (1.5), assuming throughout the paper that A in (1.5) has the form (1.10), preserving the old notations a_{ij} for entries of matrix A_0. This means that we

to Contents, page 1

assume that $A = \{a_{is}\}$, $i, s = 1, 2, \ldots, n$ with $a_{ns} = a_{sn} = 0$ for $s = 1, 2, \ldots, n-1$ and $a_{nn} = \lambda$, and $A_0 = \{a_{is}\}$, $i, s = 1, 2, \ldots, n-1$.

We will give criteria of the instability of a trivial solution of the system (1.5) if the matrix A of linear terms is defined by (1.10).

2. Preliminaries

In this part we collect the necessary material-the definition of a cone, auxiliary Chetaev-type results on instability in a cone and, finally, a third degree polynomial inequality, which will be used to estimate the sign of the full derivative of a Chetaev-type function along the trajectories of system (1.5).

2.1. Instability of the Zero Solution of Systems of Differential Equations in a Cone

We consider an autonomous system of differential equations

$$\dot{x} = f(x), \tag{2.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ satisfies a local Lipschitz condition and $f(0) = 0$, that is, (2.1) admits the trivial solution. We will consider solutions of (2.1) determined by points $(x, t) = (x_0, 0)$ where $x_0 \in \mathbb{R}^n$. The symbol $x(x_0, t)$ denotes the solution $x = x(t)$ of (2.1), satisfying initial condition $x(0) = x_0$.

Definition 2.1. The zero solution $x \equiv 0$ of (2.1) is called unstable if there exists $\varepsilon > 0$ such that, for arbitrary $\delta > 0$, there exists an $x_0 \in \mathbb{R}^n$ with $\|x_0\| < \delta$ and $T \geq 0$ such that $\|x(x_0, T)\| \geq \varepsilon$.

Definition 2.2. A set $K \subset R^n$ is called a cone if $\alpha x \in K$ for arbitrary $x \in K$ and $\alpha > 0$.

Definition 2.3. A cone K is said to be a global cone of instability for (2.1) if $x(x_0, t) \in K$ for arbitrary $x_0 \in K$ and $t \geq 0$ and $\lim_{t \to \infty} \|x(x_0, t)\| = \infty$.

Definition 2.4. The zero solution $x \equiv 0$ of (2.1) is said to be globally unstable in a cone K if K is a global cone of instability for (2.1).

Now, we prove results analogous to the classical Chetaev theorem (see, e.g., [2]) on instability in a form suitable for our analysis. As usual, if $\mathcal{S}$ is a set, then $\partial \mathcal{S}$ denotes its boundary and $\overline{\mathcal{S}}$ its closure, that is, $\overline{\mathcal{S}} := \mathcal{S} \cup \partial \mathcal{S}$.

Theorem 2.5. *Let $V : \mathbb{R}^n \to \mathbb{R}$, $V(0, \ldots, 0) = 0$ be a continuously differentiable function. Assume that the set*

$$K = \{x \in R^n : V(x) > 0\} \tag{2.2}$$

is a cone. If the full derivative of V along the trajectories of (2.1) *is positive for every $x \in K$, that is, if*

$$\dot{V}(x) := grad^T V(x) f(x) > 0, \quad x \in K, \tag{2.3}$$

then K is a global cone of instability for the system (2.1).

to Contents, page 1

Proof. Let ε be a positive number. We define a neighborhood of the origin

$$U_\varepsilon := \{x \in R^n : \|x\| < \varepsilon\}, \tag{2.4}$$

and a constant

$$M_\varepsilon := \max_{x \in \overline{U}_\varepsilon \cap \overline{K}} V(x). \tag{2.5}$$

Moreover, define a set

$$W_\delta := \left\{x \in \overline{U}_\varepsilon \cap \overline{K}, V(x) \geq \delta\right\}, \tag{2.6}$$

where δ is a positive number such that $\delta < M_\varepsilon$. Then, $W_\delta \neq \emptyset$.

Let $x_0 \in W_\delta \cap K$, then $V(x_0) = \delta_1 \in [\delta, M_\varepsilon]$. We show that there exists a $t = t_T = t_T(\varepsilon, x_0)$ such that $x(x_0, t_T) \notin \overline{U}_\varepsilon$ and $x(x_0, t_T) \in K$.

Suppose to the contrary that this is not true and $x(x_0, t) \in \overline{U}_\varepsilon$ for all $t \geq 0$. Since $\dot{V}(x) > 0$, the function V is increasing along the solutions of (2.1). Thus $x(x_0, t)$ remains in K. Due to the compactness of W_δ, there exists a positive value β such that for $x(x_0, t) \in W_\delta$

$$\frac{d}{dt} V(x(x_0, t)) = \operatorname{grad}^T V(x(x_0, t)) f(x(x_0, t)) > \beta. \tag{2.7}$$

Integrating this inequality over the interval $[0, t]$, we get

$$V(x(x_0, t)) - V(x_0) = V(x(x_0, t)) - \delta_1 > \beta t. \tag{2.8}$$

Then there exists a $t = t_T = t_T(\varepsilon, x_0)$ satisfying

$$t_T > \frac{(M_\varepsilon - \delta_1)}{\beta}, \tag{2.9}$$

such that $V(x(x_0, t_T)) > M_\varepsilon$ and, consequently, $x(x_0, t_T) \notin \overline{U}_\varepsilon$. This is contrary to our supposition. Since $\varepsilon > 0$ is arbitrary, we have

$$\lim_{t \to \infty} \|x(x_0, t)\| = \infty, \tag{2.10}$$

that is, the zero solution is globally unstable, and K is a global cone of instability. □

Theorem 2.6. *Let $V : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function and let $S, Z : \mathbb{R}^n \to \mathbb{R}$, $Z(0, \ldots, 0) = 0$ be continuous functions such that $V = S \cdot Z$. Assume that the set*

$$K_1 = \{x \in R^n : Z(x) > 0\} \tag{2.11}$$

to Contents, page 1

is a cone, and $S(x) > 0$ *for any* $x \in K_1$*. If the full derivative* (2.3) *of* V *along the trajectories of* (2.1) *is positive for every* $x \in K_1$*, that is, if* $\dot{V}(x) > 0$ *for every* $x \in K_1$*, then* K_1 *is a global cone of instability for the system* (2.1)*.*

Proof. The proof is a modification of the proof of Theorem 2.5. Let ε be a positive number. We define a neighborhood U_ε of the origin by formula (2.4) and a constant

$$M_\varepsilon := \max_{x \in \overline{U}_\varepsilon \cap \overline{K_1}} V(x). \tag{2.12}$$

Moreover, define a set

$$W_\delta := \left\{ x \in \overline{U}_\varepsilon \cap \overline{K_1}, V(x) \geq \delta \right\}, \tag{2.13}$$

where δ is a positive number such that $\delta < M_\varepsilon$. Then $W_\delta \neq \emptyset$.

Let $x_0 \in W_\delta \cap K_1$. Then $V(x_0) = \delta_1 \in [\delta, M_\varepsilon]$. We show that there exists a $t = t_T = t_T(\varepsilon, x_0)$ such that $x(x_0, t_T) \notin \overline{U}_\varepsilon$ and $x(x_0, t_T) \in K_1$.

Suppose to the contrary that this is not true and $x(x_0, t) \in \overline{U}_\varepsilon$ for all $t \geq 0$. Since $\dot{V}(x) > 0$, the function V is increasing along the solutions of (2.1). Due to the compactness of W_δ, there exists a positive value β such that for $x(x_0, t) \in W_\delta$

$$\frac{d}{dt} V(x(x_0, t)) = \operatorname{grad}^T V(x(x_0, t)) f(x(x_0, t)) > \beta. \tag{2.14}$$

Integrating this inequality over interval $[0, t]$, we get

$$V(x(x_0, t)) - V(x_0) = V(x(x_0, t)) - \delta_1 = S(x(x_0, t)) Z(x(x_0, t)) - \delta_1 > \beta t. \tag{2.15}$$

Since $S(x(x_0, t)) > 0$, the inequality

$$Z(x(x_0, t)) > \frac{\delta_1 + \beta t}{S(x(x_0, t))} > 0 \tag{2.16}$$

is an easy consequence of (2.15). Thus $x(x_0, t)$ remains in K_1. Apart from this, (2.15) also implies the existence of a $t = t_T = t_T(\varepsilon, x_0)$ satisfying

$$t_T > \frac{(M_\varepsilon - \delta_1)}{\beta}, \tag{2.17}$$

such that $V(x(x_0, t_T)) > M_\varepsilon$. Consequently, $x(x_0, t_T) \notin \overline{U}_\varepsilon$. This is contrary to our supposition. Since $\varepsilon > 0$ is arbitrary, we have

$$\lim_{t \to \infty} \|x(x_0, t)\| = \infty, \tag{2.18}$$

that is, the zero solution is globally unstable and K_1 is a global cone of instability. □

to Contents, page 1

Definition 2.7. A function V satisfying all the properties indicated in Theorem 2.5 is called a Chetaev function for the system (2.1). A function V satisfying all the properties indicated in Theorem 2.6 is called a Chetaev-type function for the system (2.1).

2.2. Auxiliary Inequality

Our results will be formulated in terms of global cones of instability. These will be derived using an auxiliary inequality valid in a given cone. Let $(x, y) \in \mathbb{R}^2$ and let k be a positive number. We define a cone

$$\mathcal{K} := \left\{(x, y) \in \mathbb{R}^2 : y > k|x|\right\}. \tag{2.19}$$

Lemma 2.8. *Let a, b, c, d, and k be given constants such that $b > 0$, $d > 0$, $k > 0$, and $|c| \leq kd$. Assume, moreover, either*

$$|a| \leq kb, \tag{2.20}$$

or

$$|a| > kb, \tag{2.21}$$

$$|c| \neq kd, \quad k \geq \max\left\{\sqrt{\frac{|a+kb|}{c+kd}}, \sqrt{\frac{|a-kb|}{|c-kd|}}\right\}, \tag{2.22}$$

then

$$ax^3 + bx^2y + cxy^2 + dy^3 > 0, \tag{2.23}$$

for every $(x, y) \in \mathcal{K}$.

Proof. We partition $\mathcal{K}$ into two disjoint cones

$$\begin{aligned} \mathcal{K}_1 &:= \left\{(x, y) \in \mathbb{R}^2 : y > k|x|,\ x > 0\right\}, \\ \mathcal{K}_2 &:= \left\{(x, y) \in \mathbb{R}^2 : y > k|x|,\ x \leq 0\right\}, \end{aligned} \tag{2.24}$$

and rewrite (2.23) as

$$x\left(ax^2 + cy^2\right) + y\left(bx^2 + dy^2\right) > 0. \tag{2.25}$$

We prove the validity of (2.23) in each of the two cones separately.

to Contents, page 1

The case of the cone $\mathcal{K}_1$. Suppose that (2.20) holds. Estimating the left-hand side of (2.25), we get

$$\begin{aligned} x\left(ax^2+cy^2\right)+y\left(bx^2+dy^2\right) &> x\left(ax^2+cy^2\right)+kx\left(bx^2+dy^2\right) \\ &= x\left[x^2(a+kb)+y^2(c+kd)\right] > 0, \end{aligned} \tag{2.26}$$

and (2.23) holds.

If inequalities (2.21) and (2.22) are valid, then, estimating the left-hand side of (2.25), we get

$$\begin{aligned} x\left(ax^2+cy^2\right)+y\left(bx^2+dy^2\right) &> x\left(ax^2+cy^2\right)+kx\left(bx^2+dy^2\right) \\ &= x\left[x^2(a+kb)+y^2(c+kd)\right] \\ &\geq x\left[-|a+kb|x^2+(c+kd)y^2\right] \\ &= (c+kd)x\left[y^2-\frac{|a+kb|}{c+kd}x^2\right] \\ &= (c+kd)x\left[y-\sqrt{\frac{|a+kb|}{c+kd}}x\right]\left[y+\sqrt{\frac{|a+kb|}{c+kd}}x\right] \\ &= (c+kd)x^2\left[k-\sqrt{\frac{|a+kb|}{c+kd}}\right]\left[k+\sqrt{\frac{|a+kb|}{c+kd}}\right] \\ &\geq 0, \end{aligned} \tag{2.27}$$

and (2.23) holds again.

The case of the cone $\mathcal{K}_2$. Suppose that (2.20) hold, then, estimating the left-hand side of (2.25), we get

$$\begin{aligned} x\left(ax^2+cy^2\right)+y\left(bx^2+dy^2\right) &= -|x|\left(ax^2+cy^2\right)+y\left(bx^2+dy^2\right) \\ &> -|x|\left(ax^2+cy^2\right)+k|x|\left(bx^2+dy^2\right) \\ &= -|x|\left[(a-kb)x^2+(c-kd)y^2\right] \\ &\geq 0, \end{aligned} \tag{2.28}$$

and (2.23) holds.

to Contents, page 1

If inequalities (2.21) and (2.22) are valid, then the estimation of (2.25) implies (we use (2.28))

$$\begin{aligned}
&x\left(ax^2+cy^2\right)+y\left(bx^2+dy^2\right)\\
&\quad> -|x|\left[(a-kb)x^2+(c-kd)y^2\right]\\
&\quad= |c-kd||x|\left[y^2-\frac{a-kb}{|c-kd|}x^2\right]\\
&\quad= \begin{cases} \geq 0 & \text{if } a-kb<0,\\[2mm] |c-kd||x|\left[y-\sqrt{\dfrac{a-kb}{|c-kd|}}\,x\right]\left[y+\sqrt{\dfrac{a-kb}{|c-kd|}}\,x\right] & \\[4mm] \geq |c-kd|x^2\left[k+\sqrt{\dfrac{a-kb}{|c-kd|}}\,\right]\left[k-\sqrt{\dfrac{a-kb}{|c-kd|}}\,\right]\geq 0 & \text{if } a-kb>0. \end{cases}
\end{aligned} \tag{2.29}$$

Hence, (2.23) holds again. □

3. Global Cone of Instability

In this part we derive a result on the instability of system (1.5) in a cone. In order to properly formulate the results, we have to define some auxiliary vectors and matrices (some definitions copy the previous ones used in Introduction, but with a dimension of $n-1$ rather than n). We denote

$$\begin{gathered}
x_{(n-1)}=(x_1,x_2,\ldots,x_{n-1})^T,\\
b_i=\left(b^i_{1n},b^i_{2n},\ldots,b^i_{n-1,n}\right)^T,\quad i=1,2,\ldots,n,\\
\tilde{b}=\left(b^1_{nn},b^2_{nn},\ldots,b^{n-1}_{nn}\right)^T.
\end{gathered} \tag{3.1}$$

Apart from this, we define symmetric $(n-1)\times(n-1)$ matrices

$$B^0_i=\left\{b^i_{sq}\right\},\quad i=1,2,\ldots,n,\quad s,q=1,2,\ldots,n-1, \tag{3.2}$$

to Contents, page 1

that is,

$$B_i^0 = \begin{pmatrix} b_{11}^i & b_{12}^i & \cdots & b_{1,n-1}^i \\ b_{21}^i & b_{22}^i & \cdots & b_{2,n-1}^i \\ \cdots & \cdots & \cdots & \cdots \\ b_{n-1,1}^i & b_{n-1,2}^i & \cdots & b_{n-1,n-1}^i \end{pmatrix},$$
$$\widetilde{B} = \begin{pmatrix} b_{1n}^1 & b_{2n}^1 & \cdots & b_{n-1,n}^1 \\ b_{1n}^2 & b_{2n}^2 & \cdots & b_{n-1,n}^2 \\ \cdots & \cdots & \cdots & \cdots \\ b_{1n}^{n-1} & b_{2n}^{n-1} & \cdots & b_{n-1,n}^{n-1} \end{pmatrix}. \tag{3.3}$$

Finally, we define an $(n-1) \times (n-1)^2$ matrix

$$\overline{B}^T = \left\{\overline{B}_1^T, \overline{B}_2^T, \ldots, \overline{B}_{n-1}^T\right\}, \tag{3.4}$$

where $(n-1) \times (n-1)$ matrices $\overline{B}_i^T$, $i = 1, 2, \ldots, n-1$ are defined as

$$\overline{B}_i^T = \begin{pmatrix} b_{i1}^1 & b_{i2}^1 & \cdots & b_{i,n-1}^1 \\ b_{i1}^2 & b_{i2}^2 & \cdots & b_{i,n-1}^2 \\ \cdots & \cdots & \cdots & \cdots \\ b_{i1}^{n-1} & b_{i2}^{n-1} & \cdots & b_{i,n-1}^{n-1} \end{pmatrix}. \tag{3.5}$$

We consider a matrix equation

$$A_0^T H + H A_0 = -C, \tag{3.6}$$

where H and C are $(n-1) \times (n-1)$ matrices. It is well-known (see, e.g., [14]) that, for a given positive definite symmetric matrix C, (3.6) can be solved for a positive definite symmetric matrix H if and only if the matrix A_0 is asymptotically stable.

Theorem 3.1 (Main result). *Assume that the matrix A_0 is asymptotically stable, $b_{nn}^n > 0$ and h is a positive number. Let C be an $(n-1) \times (n-1)$ positive definite symmetric matrix and H be a related $(n-1) \times (n-1)$ positive definite symmetric matrix solving equation* (3.6)*. Assume that the matrix $(-H\widetilde{B}^T - \widetilde{B}H + h(B_n^0)^T)$ is positive definite,*

$$\left\|2hb_n - H\widetilde{b}\right\| \le \sqrt{\lambda_{\min}(H)h} \cdot b_{nn}^n, \tag{3.7}$$

to Contents, page 1

and, in addition, one of the following conditions is valid:
either

$$\left\| H\overline{B}^T \right\| \le \sqrt{\frac{\lambda_{\min}(H)}{h}} \cdot \lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h\left(B_n^0\right)^T \right) \tag{3.8}$$

or

$$\left\| H\overline{B}^T \right\| > \sqrt{\frac{\lambda_{\min}(H)}{h}} \cdot \lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h\left(B_n^0\right)^T \right), \tag{3.9}$$

a strong inequality holds in (3.7)*, and*

$$\sqrt{\frac{\lambda_{\min}(H)}{h}} \ge \max\left\{ \sqrt{\tau_1}, \sqrt{\tau_2} \right\}, \tag{3.10}$$

where

$$\begin{aligned} \tau_1 &= \frac{\left\| H\overline{B}^T \right\| - \sqrt{\lambda_{\min}(H)/h} \cdot \lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h(B_n^0)^T \right)}{-\left\| 2hb_n - H\widetilde{b} \right\| + \sqrt{\lambda_{\min}(H)h} \cdot b_{nn}^n}, \\ \tau_2 &= \frac{\left\| H\overline{B}^T \right\| + \sqrt{\lambda_{\min}(H)/h} \cdot \lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h(B_n^0)^T \right)}{\left\| 2hb_n - H\widetilde{b} \right\| + \sqrt{\lambda_{\min}(H)h} \cdot b_{nn}^n}. \end{aligned} \tag{3.11}$$

Then the set

$$K := \left\{ \left(x_{(n-1)}^T, x_n \right) : \sqrt{h}x_n > \sqrt{x_{(n-1)}^T H x_{(n-1)}} \right\} \tag{3.12}$$

is a global cone of instability for the system (1.5)*.*

Proof. First we make auxiliary computations. For the reader's convenience, we recall that, for two $(n-1) \times (n-1)$ matrices $\mathcal{A}$, $\mathcal{A}_1$, two $1 \times (n-1)$ vectors ℓ, ℓ_1, two $(n-1) \times 1$ vectors $\mathcal{C}$, $\mathcal{C}_1$ and two 1×1 "matrices" m, m_1, the multiplicative rule

$$\begin{pmatrix} \mathcal{A} & \mathcal{C} \\ \ell & m \end{pmatrix} \begin{pmatrix} \mathcal{A}_1 & \mathcal{C}_1 \\ \ell_1 & m_1 \end{pmatrix} = \begin{pmatrix} \mathcal{A}\mathcal{A}_1 + \mathcal{C}\ell_1 & A\mathcal{C}_1 + \mathcal{C}m_1 \\ \ell\mathcal{A}_1 + m\ell_1 & \ell\mathcal{C}_1 + mm_1 \end{pmatrix} \tag{3.13}$$

holds. This rule can be modified easily for the case of arbitrary rectangular matrices under the condition that all the products are well defined.

to Contents, page 1

We will rewrite system (1.5) in an equivalent form, suitable for further investigation. With this in mind, we define an $(n-1)^2 \times (n-1)$ matrix $X_{(n-1)}$ as

$$X_{(n-1)}^T = \left(X_{1(n-1)}^T, X_{2(n-1)}^T, \ldots, X_{n-1(n-1)}^T\right), \tag{3.14}$$

where all the elements of the $(n-1) \times (n-1)$ matrices $X_{i(n-1)}^T$, $i = 1, 2, \ldots, n-1$ are equal to zero except the ith row, which equals $x_{(n-1)}^T$, that is,

$$X_{i(n-1)}^T = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \\ x_1 & x_2 & \cdots & x_{n-1} \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{3.15}$$

Moreover, we define $1 \times (n-1)$ vectors Y_i, $i = 1, 2, \ldots, n-1$ with components equal to zero except the ith element, which equals x_n, that is,

$$Y_i = (0, \ldots, 0, x_n, 0, \ldots, 0), \tag{3.16}$$

and $(n-1) \times (n-1)$ zero matrix Θ.

It is easy to see that matrices X^T and B in (1.5) can be expressed as

$$X^T = \begin{pmatrix} X_{1(n-1)}^T & Y_1^T & \cdots & X_{n-1(n-1)}^T & Y_{n-1}^T & \Theta & \theta \\ \theta^T & 0 & \cdots & \theta^T & 0 & x_{(n-1)}^T & x_n \end{pmatrix},$$

$$B = \begin{pmatrix} B_1^0 & b_1 \\ b_1^T & b_{nn}^1 \\ \cdots & \cdots \\ B_n^0 & b_n \\ b_n^T & b_{nn}^n \end{pmatrix}. \tag{3.17}$$

to Contents, page 1

Now we are able to rewrite the system (1.5) under the above assumption regarding the representation of the matrix A in the form (1.10) in an equivalent form

$$\begin{aligned} \begin{pmatrix} \dot{x}_{(n-1)} \\ \dot{x}_n \end{pmatrix} &= \begin{pmatrix} A_0 & \theta \\ \theta^T & \lambda \end{pmatrix} \begin{pmatrix} x_{(n-1)} \\ x_n \end{pmatrix} \\ &+ \begin{pmatrix} X^T_{1(n-1)} & Y^T_1 & \cdots & X^T_{n-1(n-1)} & Y^T_{n-1} & \Theta & \theta \\ \theta^T & 0 & \cdots & \theta^T & 0 & x^T_{(n-1)} & x_n \end{pmatrix} \\ &\times \begin{pmatrix} B^0_1 & b_1 \\ b^T_1 & b^1_{nn} \\ \cdots & \cdots \\ B^0_n & b_n \\ b^T_n & b^n_{nn} \end{pmatrix} \begin{pmatrix} x_{(n-1)} \\ x_n \end{pmatrix}. \end{aligned} \tag{3.18}$$

Finally, since the equalities

$$\begin{aligned} \sum_{j=1}^{n-1} X^T_{j(n-1)} B^0_j &= \overline{B}^T X_{(n-1)}, \\ \sum_{j=1}^{n-1} Y^T_j b^T_j &= \widetilde{B} x_n, \\ \sum_{j=1}^{n-1} X^T_{j(n-1)} b_j &= \widetilde{B} x_{(n-1)}, \\ \sum_{j=1}^{n-1} Y^T_j b^j_{nn} &= \widetilde{b} x_n \end{aligned} \tag{3.19}$$

can be verified easily using (3.13), we have

$$\begin{pmatrix} \dot{x}_{(n-1)} \\ \dot{x}_n \end{pmatrix} = \begin{pmatrix} A_0 + r_{11}\left(x^T_{(n-1)}, x_n\right) & r_{12}\left(x^T_{(n-1)}, x_n\right) \\ r_{21}\left(x^T_{(n-1)}, x_n\right) & \lambda + r_{22}\left(x^T_{(n-1)}, x_n\right) \end{pmatrix} \begin{pmatrix} x_{(n-1)} \\ x_n \end{pmatrix}, \tag{3.20}$$

to Contents, page 1

where

$$\begin{aligned}
r_{11}\left(x_{(n-1)}^T, x_n\right) &= \sum_{j=1}^{n-1}\left[X_{j(n-1)}^T B_j^0 + Y_j^T b_j^T\right] = \overline{B}^T X_{(n-1)} + \widetilde{B} x_n, \\
r_{12}\left(x_{(n-1)}^T, x_n\right) &= \sum_{j=1}^{n-1}\left[X_{j(n-1)}^T b_j + Y_j^T b_{nn}^j\right] = \widetilde{B} x_{(n-1)} + \widetilde{b} x_n, \\
r_{21}\left(x_{(n-1)}^T, x_n\right) &= x_{(n-1)}^T B_n^0 + x_n b_n^T, \\
r_{22}\left(x_{(n-1)}^T, x_n\right) &= x_{(n-1)}^T b_n + x_n b_{nn}^n.
\end{aligned} \tag{3.21}$$

The remaining part of the proof is based on Theorem 2.6 with a Chetaev-type function $V = S \cdot Z$ and with suitable functions S and Z. Such functions we define as

$$V\left(x_{(n-1)}^T, x_n\right) = \begin{pmatrix} x_{(n-1)}^T & x_n \end{pmatrix} \begin{pmatrix} -H & \theta \\ \theta^T & h \end{pmatrix} \begin{pmatrix} x_{(n-1)} \\ x_n \end{pmatrix}, \tag{3.22}$$

that is,

$$\begin{aligned}
V\left(x_{(n-1)}^T, x_n\right) &= -x_{(n-1)}^T H x_{(n-1)} + h x_n^2, \\
S\left(x_{(n-1)}^T, x_n\right) &= \sqrt{x_{(n-1)}^T H x_{(n-1)}} + \sqrt{h} x_n, \\
Z\left(x_{(n-1)}^T, x_n\right) &= -\sqrt{x_{(n-1)}^T H x_{(n-1)}} + \sqrt{h} x_n.
\end{aligned} \tag{3.23}$$

We will verify the necessary properties. Obviously, $V = S \cdot Z$, the set

$$\begin{aligned}
K_1 :&= \left\{\left(x_{(n-1)}^T, x_n\right) \in \mathbb{R}^n : Z\left(x_{(n-1)}, x_n\right) > 0\right\} \\
&= \left\{\left(x_{(n-1)}^T, x_n\right) \in \mathbb{R}^n : \sqrt{h} x_n > \sqrt{x_{(n-1)}^T H x_{(n-1)}}\right\}
\end{aligned} \tag{3.24}$$

is a cone and $S(x_{(n-1)}^T, x_n) > 0$ for every $(x_{(n-1)}^T, x_n) \in K_1$.

to Contents, page 1

The full derivative of V (in the form (3.22)) along the trajectories of the system (1.5) (we use its transformed form (3.20)) equals

$$\begin{aligned}\dot{V}\left(x_{(n-1)}^T, x_n\right) &= \begin{pmatrix}\dot{x}_{(n-1)}^T & \dot{x}_n\end{pmatrix}\begin{pmatrix}-H & \theta \\ \theta^T & h\end{pmatrix}\begin{pmatrix}x_{(n-1)} \\ x_n\end{pmatrix} + \begin{pmatrix}x_{(n-1)}^T & x_n\end{pmatrix}\begin{pmatrix}-H & \theta \\ \theta^T & h\end{pmatrix}\begin{pmatrix}\dot{x}_{(n-1)} \\ \dot{x}_n\end{pmatrix} \\ &= \begin{pmatrix}x_{(n-1)}^T & x_n\end{pmatrix}\begin{pmatrix}A_0^T + r_{11}^T\left(x_{(n-1)}^T, x_n\right) & r_{21}^T\left(x_{(n-1)}^T, x_n\right) \\ r_{12}^T\left(x_{(n-1)}^T, x_n\right) & \lambda + r_{22}\left(x_{(n-1)}^T, x_n\right)\end{pmatrix}\begin{pmatrix}-H & \theta \\ \theta^T & h\end{pmatrix}\begin{pmatrix}x_{(n-1)} \\ x_n\end{pmatrix} \\ &\quad + \begin{pmatrix}x_{(n-1)}^T & x_n\end{pmatrix}\begin{pmatrix}-H & \theta \\ \theta^T & h\end{pmatrix}\begin{pmatrix}A_0 + r_{11}\left(x_{(n-1)}^T, x_n\right) & r_{12}\left(x_{(n-1)}^T, x_n\right) \\ r_{21}\left(x_{(n-1)}^T, x_n\right) & \lambda + r_{22}\left(x_{(n-1)}^T, x_n\right)\end{pmatrix} \\ &\quad \times \begin{pmatrix}x_{(n-1)} \\ x_n\end{pmatrix}.\end{aligned} \tag{3.25}$$

Using formula (3.13), we get

$$\dot{V}\left(x_{(n-1)}^T, x_n\right) = \begin{pmatrix}x_{(n-1)}^T & x_n\end{pmatrix}\begin{pmatrix}c_{11}\left(x_{(n-1)}^T, x_n\right) & c_{12}\left(x_{(n-1)}^T, x_n\right) \\ c_{21}\left(x_{(n-1)}^T, x_n\right) & c_{22}\left(x_{(n-1)}^T, x_n\right)\end{pmatrix}\begin{pmatrix}x_{(n-1)} \\ x_n\end{pmatrix}, \tag{3.26}$$

where

$$\begin{aligned}c_{11}\left(x_{(n-1)}^T, x_n\right) &= -\left[A_0 + r_{11}(x_{(n-1)}^T, x_n)\right]^T H - H\left[A_0 + r_{11}\left(x_{(n-1)}^T, x_n\right)\right], \\ c_{12}\left(x_{(n-1)}^T, x_n\right) &= hr_{21}^T\left(x_{(n-1)}^T, x_n\right) - Hr_{12}\left(x_{(n-1)}^T, x_n\right), \\ c_{21}\left(x_{(n-1)}^T, x_n\right) &= hr_{21}\left(x_{(n-1)}^T, x_n\right) - r_{12}^T\left(x_{(n-1)}^T, x_n\right)H = c_{12}^T\left(x_{(n-1)}^T, x_n\right), \\ c_{22}\left(x_{(n-1)}^T, x_n\right) &= 2h\left[\lambda + r_{22}\left(x_{(n-1)}^T, x_n\right)\right].\end{aligned} \tag{3.27}$$

We reduce these formulas using (3.21). Then,

$$\begin{aligned}c_{11}\left(x_{(n-1)}^T, x_n\right) &= -\left(A_0^T H + HA_0\right) - \left(\overline{B}^T X_{(n-1)} + \widetilde{B}x_n\right)^T H - H\left(\overline{B}^T X_{(n-1)} + \widetilde{B}x_n\right), \\ c_{12}\left(x_{(n-1)}^T, x_n\right) &= h\left(x_{(n-1)}^T B_n^0 + x_n b_n^T\right)^T - H\left(\widetilde{B}x_{(n-1)} + \widetilde{b}x_n\right), \\ c_{21}\left(x_{(n-1)}^T, x_n\right) &= h\left(x_{(n-1)}^T B_n^0 + x_n b_n^T\right) - \left(\widetilde{B}x_{(n-1)} + \widetilde{b}x_n\right)^T H, \\ c_{22}\left(x_{(n-1)}^T, x_n\right) &= 2h\left(\lambda + x_{(n-1)}^T b_n + x_n b_{nn}^n\right).\end{aligned} \tag{3.28}$$

The derivative (3.26) turns into

$$\begin{aligned}
\dot{V}\left(x_{(n-1)}^T, x_n\right) &= x_{(n-1)}^T c_{11}\left(x_{(n-1)}^T, x_n\right)x_{(n-1)} + x_{(n-1)}^T c_{12}\left(x_{(n-1)}^T, x_n\right)x_n \\
&\quad + x_n c_{21}\left(x_{(n-1)}^T, x_n\right)x_{(n-1)} + x_n c_{22}\left(x_{(n-1)}^T, x_n\right)x_n \\
&= x_{(n-1)}^T\left[-\left(A_0^T H + HA_0\right) - \left(\overline{B}^T X_{(n-1)} + \widetilde{B}x_n\right)^T H - H\left(\overline{B}^T X_{(n-1)} + \widetilde{B}x_n\right)\right]x_{(n-1)} \\
&\quad + x_{(n-1)}^T\left[h\left(x_{(n-1)}^T B_n^0 + x_n b_n^T\right)^T - H\left(\widetilde{B}x_{(n-1)} + \widetilde{b}x_n\right)\right]x_n \\
&\quad + x_n\left[h\left(x_{(n-1)}^T B_n^0 + x_n b_n^T\right) - \left(\widetilde{B}x_{(n-1)} + \widetilde{b}x_n\right)^T H\right]x_{(n-1)} \\
&\quad + x_n\left[2h\left(\lambda + x_{(n-1)}^T b_n + x_n b_{nn}^n\right)\right]x_n \\
&= -x_{(n-1)}^T\left(A_0^T H + HA_0\right)x_{(n-1)} + 2h\lambda x_n^2 \\
&\quad - x_{(n-1)}^T\left(\left(\overline{B}^T X_{(n-1)}\right)^T H + H\overline{B}^T X_{(n-1)}\right)x_{(n-1)} \\
&\quad - x_{(n-1)}^T\left(\left(\widetilde{B}x_n\right)^T H + H\widetilde{B}x_n\right)x_{(n-1)} \\
&\quad + x_{(n-1)}^T\left(2h\left(B_n^0\right)^T - H\widetilde{B} - \widetilde{B}H\right)x_{(n-1)}x_n \\
&\quad + 2x_{(n-1)}^T\left(hb_n - H\widetilde{b}\right)x_n^2 \\
&\quad + 2h\left(x_{(n-1)}^T b_n + x_n b_{nn}^n\right)x_n^2.
\end{aligned} \tag{3.29}$$

Finally, using (3.6), we get

$$\begin{aligned}
\dot{V}\left(x_{(n-1)}^T, x_n\right) &= x_{(n-1)}^T C x_{(n-1)} + 2h\lambda x_n^2 - 2x_{(n-1)}^T H\overline{B}^T X_{(n-1)}x_{(n-1)} \\
&\quad + 2x_{(n-1)}^T\left[-H\widetilde{B}^T - \widetilde{B}H + h\left(B_n^0\right)^T\right]x_{(n-1)}x_n + 2x_{(n-1)}^T\left(2hb_n - H\widetilde{b}\right)x_n^2 + 2hb_{nn}^n x_n^3.
\end{aligned} \tag{3.30}$$

to Contents, page 1

Let us find the conditions for the positivity of $\dot{V}(x^T_{(n-1)}, x_n)$ in the cone K_1. We use (3.30). If $(x^T_{(n-1)}, x_n) \in K_1$, then $x_n \geq 0$ and

$$\begin{aligned} \dot{V}\left(x^T_{(n-1)}, x_n\right) \geq{}& x^T_{(n-1)} C x_{(n-1)} + 2h\lambda x_n^2 - 2\left\| H\overline{B}^T \right\| \cdot \left\| x_{(n-1)} \right\|^3 \\ &+ 2\lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h\left(B_n^0\right)^T\right) \cdot \left\| x_{(n-1)} \right\|^2 \cdot x_n \\ &- 2\left\| 2hb_n - H\widetilde{b} \right\| \cdot \left\| x_{(n-1)} \right\| \cdot x_n^2 + 2hb_{nn}^n x_n^3. \end{aligned} \tag{3.31}$$

We set

$$\begin{aligned} a &= -2\left\| H\overline{B}^T \right\|, \\ b &= 2\lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h\left(B_n^0\right)^T\right), \\ c &= -2\left\| 2hb_n - H\widetilde{b} \right\|, \\ d &= 2hb_{nn}^n. \end{aligned} \tag{3.32}$$

If

$$a\left\| x_{(n-1)} \right\|^3 + b\left\| x_{(n-1)} \right\|^2 \cdot x_n + c\left\| x_{(n-1)} \right\| \cdot x_n^2 + dx_n^3 > 0 \tag{3.33}$$

in K_1, then $\dot{V}(x^T_{(n-1)}, x_n) > 0$ since C is a positive definite matrix and

$$x^T_{(n-1)} C x_{(n-1)} + 2h\lambda x_n^2 \geq \lambda_{\min}(C)\left\| x_{(n-1)} \right\|^2 + 2h\lambda x_n^2 > 0. \tag{3.34}$$

If $(x^T_{(n-1)}, x_n) \in K_1$, then

$$x_n > \sqrt{\frac{x^T_{(n-1)} H x_{(n-1)}}{h}} \geq \sqrt{\frac{\lambda_{\min}(H)}{h}} \cdot \left\| x_{(n-1)} \right\|, \tag{3.35}$$

$$K_1 \subset \mathcal{K}^* := \left\{ \left(x^T_{(n-1)}, x_n\right) \in \mathbb{R}^n : x_n > \sqrt{\frac{\lambda_{\min}(H)}{h}} \cdot \left\| x_{(n-1)} \right\| \right\}. \tag{3.36}$$

Now, we use Lemma 2.8 with $\mathcal{K} = \mathcal{K}^*$, $y = x_n$, $x = \|x_{(n-1)}\|$, with coefficients a, b, c, and d defined by formula (3.32) and with $k := \sqrt{\lambda_{\min}(H)/h}$.

Obviously $|c| \leq kd$ because, due to (3.7), inequality

$$\left\| 2hb_n - H\widetilde{b} \right\| \leq \sqrt{\lambda_{\min}(H)h} \cdot b_{nn}^n \tag{3.37}$$

to Contents, page 1

holds. Moreover, $|a| \le kb$ if (3.8) holds, that is, if

$$\left\|H\overline{B}^T\right\| \le \sqrt{\frac{\lambda_{\min}(H)}{h}} \cdot \lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h\left(B_n^0\right)^T\right). \tag{3.38}$$

Further, $|a| > kb$ if (3.9) holds, that is, if

$$\left\|H\overline{B}^T\right\| > \sqrt{\frac{\lambda_{\min}(H)}{h}} \cdot \lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h_n\left(B_n^0\right)^T\right), \tag{3.39}$$

and (2.22) holds due to (4.10) and the condition $|c| \ne kd$. Thus the assumptions of Lemma 2.8 are true, the inequality (3.33) holds in the cone $\mathcal{K}^*$ and, due to embedding (3.36), in the cone K_1 as well.

All the assumptions of Theorem 2.6 are fulfilled with regard to system (1.5) and the theorem is proved, because $K_1 = K$. □

Remark 3.2. We will focus our attention to Lemma 2.8 about the positivity of a third-degree polynomial in two variables in the cone $\mathcal{K}$. We used it to estimate the derivative $\dot{V}$ expressed by formula (3.30). Obviously, there are other possibilities of estimating its sign. Let us demonstrate one of them. Let us, for example, estimate the right-hand side of (3.31) in the cone K_1 using inequality (3.35), then

$$\begin{aligned}
\dot{V}\left(x_{(n-1)}^T, x_n\right) \ge{}& x_{(n-1)}^T C x_{(n-1)} + 2h\lambda x_n^2 - 2\left\|H\overline{B}^T\right\| \cdot \left\|x_{(n-1)}\right\|^3 \\
&+ 2\lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h\left(B_n^0\right)^T\right) \cdot \left\|x_{(n-1)}\right\|^2 \cdot x_n \\
&- 2\left\|2hb_n - H\widetilde{b}\right\| \cdot \left\|x_{(n-1)}\right\| \cdot x_n^2 + 2hb_{nn}^n x_n^3 \\
\ge{}& \lambda_{\min}(C)\left\|x_{(n-1)}\right\|^2 + 2h\lambda x_n^2 - 2\left\|H\overline{B}^T\right\| \cdot \left\|x_{(n-1)}\right\|^3 \\
&+ 2\sqrt{\frac{\lambda_{\min}(H)}{h}} \cdot \lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h\left(B_n^0\right)^T\right) \cdot \left\|x_{(n-1)}\right\|^3 \\
&- 2\left\|2hb_n - H\widetilde{b}\right\| \cdot \left\|x_{(n-1)}\right\| \cdot x_n^2 + 2\sqrt{\frac{\lambda_{\min}(H)}{h}} \cdot \left\|x_{(n-1)}\right\| \cdot hb_{nn}^n \cdot x_n^2,
\end{aligned} \tag{3.40}$$

and the positivity of $\dot{V}(x_{(n-1)}^T, x_n)$ will be guaranteed if

$$\begin{gathered}
\left\|H\overline{B}^T\right\| \le \sqrt{\frac{\lambda_{\min}(H)}{h}} \cdot \lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h\left(B_n^0\right)^T\right), \\
\left\|2hb_n - H\widetilde{b}\right\| \le \sqrt{\lambda_{\min}(H)h} \cdot b_{nn}^n.
\end{gathered} \tag{3.41}$$

to Contents, page 1

We see that this approach produces only one set of inequalities for the positivity of $\dot{V}(x_{(n-1)}^T, x_n)$, namely the case when (3.7) and (3.8) holds. Unfortunately, using such approach, we are not able to detect the second case (3.7) and (3.9) when $\dot{V}(x_{(n-1)}^T, x_n)$ is positive. This demonstrates the advantage of detailed estimates using the above third-degree polynomial in two variables.

4. Planar Case

Now we consider a particular case of the system (1.5) for $n = 2$. This means that, in accordance with (1.5) and (1.10), we consider a system

$$\begin{aligned}
\dot{x}_1(t) &= ax_1(t) + b_{11}^1 x_1^2(t) + 2b_{12}^1 x_1(t)x_2(t) + b_{22}^1 x_2^2(t), \\
\dot{x}_2(t) &= \lambda x_2(t) + b_{11}^2 x_1^2(t) + 2b_{12}^2 x_1(t)x_2(t) + b_{22}^2 x_2^2(t),
\end{aligned} \tag{4.1}$$

where $a < 0$ and $\lambda > 0$. The solution of matrix equation (3.6) for $A_0 = (a)$, $H = (h_{11})$, and $C = (c)$ with $c > 0$, that is,

$$(ah_{11}) + (h_{11}a) = -(c) \tag{4.2}$$

gives

$$H = (h_{11}) = \left(-\frac{c}{2a}\right), \tag{4.3}$$

with $h_{11} = -c/2a > 0$. The set K defined by (3.12) where $h > 0$ and $x_{(n-1)} = x_1$ reduces to

$$K = \left\{(x_1, x_2) : x_2 > \sqrt{\frac{c}{2|a|h}} \cdot |x_1|\right\}. \tag{4.4}$$

Now, from Theorem 3.1, we will deduce sufficient conditions indicating K being a global cone of instability for system (4.1). In our particular case, we have

$$\begin{aligned}
b_i = \left(b_{12}^i\right), \quad i = 1, 2, \quad \tilde{b} = \left(b_{22}^1\right), \\
B_i^0 = \left(b_{11}^i\right), \quad i = 1, 2, \quad \tilde{B} = (b_{12}^1), \quad \overline{B}^T = (b_{11}^1) = B_1^0.
\end{aligned} \tag{4.5}$$

to Contents, page 1

Now, we compute all necessary expressions used in Theorem 3.1. We have

$$-H\widetilde{B}^T - \widetilde{B}H + h\left(B_n^0\right)^T = -\left(-\frac{c}{2a}\right)\left(b_{12}^1\right) - \left(b_{12}^1\right)\left(-\frac{c}{2a}\right) + h\left(b_{11}^2\right) = \left(hb_{11}^2 - \frac{c}{|a|}b_{12}^1\right),$$

$$\left\|2hb_n - H\widetilde{b}\right\| = \left|2hb_{12}^2 - \frac{c}{2|a|}b_{22}^1\right|,$$

$$\sqrt{\lambda_{\min}(H)h} = \sqrt{\frac{ch}{2|a|}},$$

$$\sqrt{\frac{\lambda_{\min}(H)}{h}} = \sqrt{\frac{c}{2|a|h}},$$

$$\left\|H\overline{B}^T\right\| = \left|\frac{c}{2|a|}b_{11}^1\right| = \frac{c}{2|a|}\left|b_{11}^1\right|,$$

$$\lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h\left(B_n^0\right)^T\right) = hb_{11}^2 - \frac{c}{|a|}b_{12}^1, \tag{4.6}$$

$$\begin{aligned}
\tau_1 &= \frac{\left\|H\overline{B}^T\right\| - \sqrt{\lambda_{\min}(H)/h}\cdot\lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h(B_n^0)^T\right)}{-\left\|2hb_n - H\widetilde{b}\right\| + \sqrt{\lambda_{\min}(H)h}\cdot b_{nn}^n} \\
&= \frac{(c/2|a|)\left|b_{11}^1\right| - \sqrt{c/2|a|h}\cdot\left(hb_{11}^2 - (c/|a|)b_{12}^1\right)}{-\left|2hb_{12}^2 - (c/2|a|)b_{22}^1\right| + \sqrt{ch/2|a|}\cdot b_{22}^2},
\end{aligned}$$

$$\begin{aligned}
\tau_2 &= \frac{\left\|H\overline{B}^T\right\| + \sqrt{\lambda_{\min}(H)/h}\cdot\lambda_{\min}\left(-H\widetilde{B}^T - \widetilde{B}H + h(B_n^0)^T\right)}{\left\|2hb_n - H\widetilde{b}\right\| + \sqrt{\lambda_{\min}(H)h}\cdot b_{nn}^n} \\
&= \frac{(c/2|a|)\left|b_{11}^1\right| + \sqrt{(c/2|a|h)}\cdot\left(hb_{11}^2 - (c/|a|)b_{12}^1\right)}{\left|2hb_{12}^2 - (c/2|a|)b_{22}^1\right| + \sqrt{(ch/2|a|)}\cdot b_{22}^2}.
\end{aligned}$$

Theorem 4.1 (Planar Case). *Assume that $a < 0$, $b_{22}^2 > 0$, $h > 0$, $c > 0$ and $hb_{11}^2|a| > cb_{12}^1$. Let*

$$\left|2hb_{12}^2 - \frac{c}{2|a|}b_{22}^1\right| \le \sqrt{\frac{ch}{2|a|}}\cdot b_{22}^2, \tag{4.7}$$

and, in addition, one of the following conditions is valid:
either

$$\frac{c}{2|a|}\left|b_{11}^1\right| \le \sqrt{\frac{c}{2|a|h}}\cdot\left(hb_{11}^2 - \frac{c}{|a|}b_{12}^1\right) \tag{4.8}$$

to Contents, page 1

or

$$\frac{c}{2|a|}\left|b_{11}^{1}\right| > \sqrt{\frac{c}{2|a|h}} \cdot \left(hb_{11}^{2} - \frac{c}{|a|}b_{12}^{1}\right), \tag{4.9}$$

strong inequality holds in (4.7)*, and*

$$\sqrt{\frac{c}{2|a|h}} \geq \max\left\{\sqrt{\tau_1}, \sqrt{\tau_2}\right\}, \tag{4.10}$$

where τ_1 *and* τ_1 *are defined by* (4.6)*. Then the set* K *defined by* (4.4) *is a global cone of instability for the system* (4.1)*.*

It is easy to see that the choice $h = 1$, $c = |a|$ significantly simplifies all assumptions. Therefore we give such a particular case of Theorem 4.1.

Corollary 4.2 (Planar Case). *Assume that* $a < 0$, $b_{22}^{2} > 0$ *and* $b_{11}^{2} > b_{12}^{1}$*. Let*

$$\left|2b_{12}^{2} - \frac{1}{2}b_{22}^{1}\right| \leq \frac{1}{\sqrt{2}} \cdot b_{22}^{2}, \tag{4.11}$$

and, in addition, one of the following conditions is valid:

either

$$\frac{1}{2}\left|b_{11}^{1}\right| \leq \frac{1}{\sqrt{2}} \cdot \left(b_{11}^{2} - b_{12}^{1}\right) \tag{4.12}$$

or

$$\frac{1}{2}\left|b_{11}^{1}\right| > \frac{1}{\sqrt{2}} \cdot \left(b_{11}^{2} - b_{12}^{1}\right), \tag{4.13}$$

strong inequality holds in (4.11)*, and*

$$\frac{1}{\sqrt{2}} \geq \max\left\{\sqrt{\tau_1}, \sqrt{\tau_2}\right\}, \tag{4.14}$$

where

$$\tau_1 = \frac{(1/2)\left|b_{11}^{1}\right| - \left(1/\sqrt{2}\right) \cdot \left(b_{11}^{2} - b_{12}^{1}\right)}{-\left|2b_{12}^{2} - (1/2)b_{22}^{1}\right| + \left(1/\sqrt{2}\right) \cdot b_{22}^{2}}, \qquad \tau_2 = \frac{(1/2)\left|b_{11}^{1}\right| + \left(1/\sqrt{2}\right) \cdot \left(b_{11}^{2} - b_{12}^{1}\right)}{\left|2b_{12}^{2} - (1/2)b_{22}^{1}\right| + \left(1/\sqrt{2}\right) \cdot b_{22}^{2}}, \tag{4.15}$$

to Contents, page 1

Then the set

$$K = \left\{ (x_1, x_2) : x_2 > \frac{1}{\sqrt{2}} \cdot |x_1| \right\} \tag{4.16}$$

is a global cone of instability for the system (4.1).

Example 4.3. The set K defined by (4.16) is a global cone of instability for the system

$$\begin{aligned} \dot{x}_1(t) &= ax_1(t) + x_1^2(t) + 2\sqrt{2}x_1(t)x_2(t) + x_2^2(t), \\ \dot{x}_2(t) &= \lambda x_2(t) + 2\sqrt{2}x_1^2(t) + 2x_1(t)x_2(t) + 2\sqrt{2}x_2^2(t), \end{aligned} \tag{4.17}$$

where $a < 0$ and $\lambda > 0$ since inequalities (4.11) and (4.12) in Corollary 4.2 hold.

Example 4.4. The set K defined by (4.16) is a global cone of instability for the system

$$\begin{aligned} \dot{x}_1(t) &= ax_1(t) + 4x_1^2(t) + 2\sqrt{2}x_1(t)x_2(t) + x_2^2(t), \\ \dot{x}_2(t) &= \lambda x_2(t) + 2\sqrt{2}x_1^2(t) + 2x_1(t)x_2(t) + 20\sqrt{2}x_2^2(t), \end{aligned} \tag{4.18}$$

where $a < 0$ and $\lambda > 0$ since inequalities (4.11), (4.13), (4.14) in Corollary 4.2 hold.

Acknowledgments

This research was supported by Grants nos. P201/11/0768 and P201/10/1032 of Czech Grant Agency, and by the Council of Czech Government nos. MSM 0021630503, MSM 0021630519, and MSM 0021630529, and by Grant FEKT-S-11-2-921 of Faculty of Electrical Engineering and Communication.

References

[1] A. M. Lyapunov, *General Problem of Stability of Motion*, Gostekhizdat, Moscow, Russia, 1950.
[2] N. G. Chetaev, *Dynamic Stability*, Nauka, Moscow, Russia, 1965.
[3] B. P. Demidovič, *Lectures on Mathematical Stability Theory*, Nauka, Moscow, Russia, 1967.
[4] X. Liao, L. Wang, and P. Yu, *Stability of Dynamical Systems*, vol. 5 of *Monograph Series on Nonlinear Science and Complexity*, Elsevier, Amsterdam, The Netherlands, 2007.
[5] V. M. Matrosov and A. I. Malikov, "The development of the ideas of A. M. Lyapunov over one hundred years: 1892–1992," *Izvestiya Vysshikh Uchebnykh Zavedeniĭ. Matematika*, no. 4, pp. 3–47, 1993 (Russian).
[6] K. P. Persidskiy, "About dynamic stability in the first approximation," *Matematicheskii Sbornik*, vol. 40, pp. 284–293, 1933 (Russian).
[7] S. K. Persidskiy, "Investigation by the second Lyapunov method of the property of having fixed sign in a cone of homogeneous polynomials of the second and third power," *Dynam. Sys.*, vol. 33, pp. 3–6, 1994.
[8] N. Rouche, P. Habets, and M. Laloy, *Stability Theory by Liapunov's Method*, vol. 2 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 1977.
[9] V. V. Rumyantsev, "Development of research in the USSR in the theory of stability of motion," *Differential Equations*, vol. 19, no. 5, pp. 739–776, 1983 (Russian).

to Contents, page 1

[10] V. E. Slyusarchuk, *Instability of Solutions of Evolutionary Equations*, NUVGP, Rovno, Ukraine, 2004.
[11] F. P. Gantmacher, *The Theory of Matrices*, vol. 1, AMS Chelsea, Providence, RI, USA, 2002.
[12] D. I. Martynjuk, *Lectures on the Qualitative Theory of Difference Equations*, Naukova Dumka, Kiev, Ukraine, 1972.
[13] J. Diblík and D. Khusainov, "Stability of an autonomous system with quadratic right-hand side in the critical case," *Nonlinear Dynamics and Systems Theory*, vol. 2, no. 2, pp. 145–156, 2002.
[14] T. A. Burton, *Stability and Periodic Solutions of Ordinary and Functional Differential Equations*, Dover, Mineola, NY, USA, 2005.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 970469, 14 pages
doi:10.1155/2011/970469

Research Article

Invariant Sets of Impulsive Differential Equations with Particularities in ω-Limit Set

Mykola Perestyuk and Petro Feketa

Faculty of Mechanics and Mathematics, Taras Shevchenko National University of Kyiv, Volodymyrska Street 64, 01033 Kyiv, Ukraine

Correspondence should be addressed to Petro Feketa, petro.feketa@gmail.com

Received 19 January 2011; Accepted 14 February 2011

Academic Editor: Elena Braverman

Sufficient conditions for the existence and asymptotic stability of the invariant sets of an impulsive system of differential equations defined in the direct product of a torus and an Euclidean space are obtained.

1. Introduction

The evolution of variety of processes in physics, chemistry, biology, and so forth, frequently undergoes short-term perturbations. It is sometimes convenient to neglect the duration of the perturbations and to consider these perturbations to be "instantaneous." This leads to the necessity of studying the differential equations with discontinuous trajectories, the so-called impulsive differential equations. The fundamentals of the mathematical theory of impulsive differential equations are stated in [1–4]. The theory is developing intensively due to its applied value in simulations of the real world phenomena.

At the same time, this paper is closely related to the oscillation theory. In the middle of the 20th century, a sharp turn towards the investigations of the oscillating processes that were characterized as "almost exact" iterations within "almost the same" periods of time took place. Quasiperiodic oscillations were brought to the primary focus of investigations of the oscillation theory [5].

Quasiperiodic oscillations are a sufficiently complicated and sensitive object for investigating. The practical value of indicating such oscillations is unessential. Due to the instability of frequency basis, quasiperiodic oscillation collapses easily and may be transformed into periodic oscillation via small shift of the right-hand side of the system. This fact has led to search for more rough object than the quasiperiodic solution. Thus the minimal set that is covered by the trajectories of the quasiperiodic motions becomes the main object of investigations. As it is known, such set is a torus. The first profound assertions regarding the

to Contents, page 1

invariant toroidal manifolds were obtained by Bogoliubov et al. [6, 7]. Further results in this area were widely extended by many authors.

Consider the system of differential equations

$$\frac{dz}{dt} = F(z), \tag{1.1}$$

where the function $F(z)$ is defined in some subset D of the $(m+n)$-dimensional Euclidean space E^{m+n}, continuous and satisfies a Lipschitz condition. Let M be an invariant toroidal manifold of the system. While investigating the trajectories that begin in the neighborhood of the manifold M, it is convenient to make the change of variables from Euclidean coordinates $(z_1, \ldots, z_{m+n})$ to so-called local coordinates $\varphi = (\varphi_1, \ldots, \varphi_m)$, $x = (x_1, \ldots, x_n)$, where φ is a point on the surface of an m-dimensional torus $\mathcal{T}^m$ and x is a point in an n-dimensional Euclidean space E^n. The change of variables is performed in a way such that the equation, which defines the invariant manifold M, transforms into $x = 0$, $\varphi \in \mathcal{T}^m$ in the new coordinates. In essence, the manifold $x = 0$, $\varphi \in \mathcal{T}^m$ is the m-dimensional torus in the space $\mathcal{T}^m \times E^n$. The character of stability of the invariant torus M is closely linked with stability of the set $x = 0$, $\varphi \in \mathcal{T}^m$: from stability, asymptotic stability, and instability of the manifold M, there follow the stability, asymptotic stability, and instability of the torus $x = 0$, $\varphi \in \mathcal{T}^m$ correspondingly and vice versa. This is what determines the relevance and value of the investigation of conditions for the existence and stability of invariant sets of the systems of differential equations defined in $\mathcal{T}^m \times E^n$. Theory of the existence and perturbation, properties of smoothness, and stability of invariant sets of systems defined in $\mathcal{T}^m \times E^n$ are considered in [8].

2. Preliminaries

The main object of investigation of this paper is the system of differential equations, defined in the direct product of an m-dimensional torus $\mathcal{T}^m$ and an n-dimensional Euclidean space E^n that undergo impulsive perturbations at the moments when the phase point φ meets a given set in the phase space. Consider the system

$$\begin{aligned} \frac{d\varphi}{dt} &= a(\varphi), \\ \frac{dx}{dt} &= A(\varphi)x + f(\varphi), \quad \varphi \notin \Gamma, \\ \Delta x|_{\varphi\in\Gamma} &= B(\varphi)x + g(\varphi), \end{aligned} \tag{2.1}$$

where $\varphi = (\varphi_1, \ldots, \varphi_m)^T \in \mathcal{T}^m$, $x = (x_1, \ldots, x_n)^T \in E^n$, $a(\varphi)$ is a continuous 2π-periodic with respect to each of the components φ_v, $v = 1, \ldots, m$ vector function that satisfies a Lipschitz condition

$$\left\| a(\varphi'') - a(\varphi') \right\| \leq L \left\| \varphi'' - \varphi' \right\| \tag{2.2}$$

for every $\varphi', \varphi'' \in \mathcal{T}^m$. $A(\varphi), B(\varphi)$ are continuous 2π-periodic with respect to each of the components φ_v, $v = 1, \ldots, m$ square matrices; $f(\varphi)$, $g(\varphi)$ are continuous (piecewise continuous with first kind discontinuities in the set Γ) 2π-periodic with respect to each of the components φ_v, $v = 1, \ldots, m$ vector functions.

to Contents, page 1

Some aspects regarding existence and stability of invariant sets of systems similar to (2.1) were considered by different authors in [9–12].

We regard the point $\varphi = (\varphi_1, \dots, \varphi_m)^T$ as a point of the m-dimensional torus $\mathcal{T}^m$ so that the domain of the functions $A(\varphi), B(\varphi), f(\varphi), g(\varphi)$, and $a(\varphi)$ is the torus $\mathcal{T}^m$. We assume that the set Γ is a subset of the torus $\mathcal{T}^m$, which is a manifold of dimension $m-1$ defined by the equation $\Phi(\varphi) = 0$ for some continuous scalar 2π-periodic with respect to each of the components φ_v, $v = 1, \dots, m$ function.

The system of differential equations

$$\frac{d\varphi}{dt} = a(\varphi) \tag{2.3}$$

defines a dynamical system on the torus $\mathcal{T}^m$. Denote by $\varphi_t(\varphi)$ the solution of (2.3) that satisfies the initial condition $\varphi_0(\varphi) = \varphi$. The Lipschitz condition (2.2) guarantees the existence and uniqueness of such solution. Moreover, the solutions $\varphi_t(\varphi)$ satisfies a group property [8]

$$\varphi_t(\varphi_\tau(\varphi)) = \varphi_{t+\tau}(\varphi) \tag{2.4}$$

for all $t, \tau \in \mathbb{R}$ and $\varphi \in \mathcal{T}^m$.

Denote by $t_i(\varphi)$, $i \in \mathbb{Z}$ the solutions of the equation $\Phi(\varphi_t(\varphi)) = 0$ that are the moments of impulsive action in system (2.1). Let the function $\Phi(\varphi)$ be such that the solutions $t = t_i(\varphi)$ exist, since otherwise, system (2.1) would not be an impulsive system. Assume that

$$\begin{gathered} \lim_{i \to \pm\infty} t_i(\varphi) = \pm\infty, \\ \lim_{T \to \pm\infty} \frac{i(t, t+T)}{T} = p < \infty \end{gathered} \tag{2.5}$$

uniformly with respect to $t \in \mathbb{R}$, where $i(a, b)$ is the number of the points $t_i(\varphi)$ in the interval (a, b). Hence, the moments of impulsive perturbations $t_i(\varphi)$ satisfy the equality [10, 11]

$$t_i(\varphi_{-t}(\varphi)) - t_i(\varphi) = t. \tag{2.6}$$

Together with system (2.1), we consider the linear system

$$\begin{gathered} \frac{dx}{dt} = A(\varphi_t(\varphi))x + f(\varphi_t(\varphi)), \quad t \neq t_i(\varphi), \\ \Delta x|_{t=t_i(\varphi)} = B(\varphi_{t_i(\varphi)}(\varphi))x + g(\varphi_{t_i(\varphi)}(\varphi)) \end{gathered} \tag{2.7}$$

that depends on $\varphi \in \mathcal{T}^m$ as a parameter. We obtain system (2.7) by substituting $\varphi_t(\varphi)$ for φ in the second and third equations of system (2.1). By invariant set of system (2.1), we understand a set that is defined by a function $u(\varphi)$, which has a period 2π with respect to each of the components φ_v, $v = 1, \dots, m$, such that the function $x(t, \varphi) = u(\varphi_t(\varphi))$ is a solution of system (2.7) for every $\varphi \in \mathcal{T}^m$.

to Contents, page 1

We call a point φ^* an ω-limit point of the trajectory $\varphi_t(\varphi)$ if there exists a sequence $\{t_n\}_{n\in\mathbb{N}}$ in $\mathbb{R}$ so that

$$\lim_{n\to+\infty} t_n = +\infty, \qquad \lim_{n\to+\infty} \varphi_{t_n}(\varphi) = \varphi^*. \tag{2.8}$$

The set of all ω-limit points for a given trajectory $\varphi_t(\varphi)$ is called ω-limit set of the trajectory $\varphi_t(\varphi)$ and denoted by Ω_φ.

Referring to system (2.7), the matrices $A(\varphi_t(\varphi))$ and $B(\varphi_t(\varphi))$, that influence the behavior of the solution $x(t,\varphi)$ of the system (2.7), depend not only on the functions $A(\varphi)$ and $B(\varphi)$ but also on the behavior of the trajectories $\varphi_t(\varphi)$. Moreover, in [9], sufficient conditions for the existence and stability of invariant sets of a system similar to (2.1) were obtained in terms of a Lyapunov function $V(\varphi,x)$ that satisfies some conditions in the domain $Z = \{\varphi \in \Omega,\ x \in \overline{J}_h\}$, where $\overline{J}_h = \{x \in E^n,\ \|x\| \le h,\ h > 0\}$,

$$\Omega = \bigcup_{\varphi\in\mathcal{T}^m} \Omega_\varphi. \tag{2.9}$$

Since the Lyapunov function has to satisfy some conditions not on the whole surface of the torus $\mathcal{T}^m$ but only in the ω-limit set Ω, it is interesting to consider system (2.1) with specific properties in the domain Ω.

3. Main Result

Consider system (2.1) assuming that the matrices $A(\varphi)$ and $B(\varphi)$ are constant in the domain Ω:

$$A(\varphi)\big|_{\varphi\in\Omega} = \widetilde{A}, \qquad B(\varphi)\big|_{\varphi\in\Omega} = \widetilde{B}. \tag{3.1}$$

Therefore, for every $\varphi \in \mathcal{T}^m$

$$\lim_{t\to+\infty} A(\varphi_t(\varphi)) = \widetilde{A}, \qquad \lim_{t\to+\infty} B(\varphi_t(\varphi)) = \widetilde{B}. \tag{3.2}$$

We will obtain sufficient conditions for the existence and asymptotic stability of an invariant set of the system (2.1) in terms of the eigenvalues of the matrices $\widetilde{A}$ and $\widetilde{B}$. Denote by

$$\gamma = \max_{j=1,\dots,n} \operatorname{Re}\lambda_j\left(\widetilde{A}\right), \qquad \alpha^2 = \max_{j=1,\dots,n} \lambda_j\left(\left(E+\widetilde{B}\right)^T\left(E+\widetilde{B}\right)\right). \tag{3.3}$$

Similar systems without impulsive perturbations have been considered in [13].

to Contents, page 1

Theorem 3.1. *Let the moments of impulsive perturbations $\{t_i(\varphi)\}$ be such that uniformly with respect to $t \in \mathbb{R}$ there exists a finite limit*

$$\lim_{\widetilde{T}\to\infty} \frac{i\left(t, t+\widetilde{T}\right)}{\widetilde{T}} = p. \tag{3.4}$$

If the following inequality holds

$$\gamma + p \ln \alpha < 0, \tag{3.5}$$

then system (2.1) *has an asymptotically stable invariant set.*

Proof. Consider a homogeneous system of differential equations

$$\begin{aligned} \frac{dx}{dt} &= A(\varphi_t(\varphi))x, \quad t \neq t_i(\varphi), \\ \Delta x|_{t=t_i(\varphi)} &= B(\varphi_{t_i(\varphi)}(\varphi))x \end{aligned} \tag{3.6}$$

that depends on $\varphi \in \mathcal{T}^m$ as a parameter. By $\Omega_\tau^t(\varphi)$, we denote the fundamental matrix of system (3.6), which turns into an identity matrix at the point $t = \tau$, that is, $\Omega_\tau^\tau(\varphi) \equiv E$. It can be readily verified [4] that $\Omega_\tau^t(\varphi)$ satisfies the equalities

$$\begin{aligned} \frac{\partial}{\partial t}\Omega_\tau^t(\varphi) &= A(\varphi_t(\varphi))\Omega_\tau^t(\varphi), \\ \Omega_\tau^t(\varphi) &= \Omega_\tau^t(\varphi + 2\pi e_k), \\ \Omega_{t+\tau}^t(\varphi_{-t}(\varphi)) &= \Omega_\tau^0(\varphi) \end{aligned} \tag{3.7}$$

for all $t, \tau \in \mathbb{R}$ and $\varphi \in \mathcal{T}^m$. Rewrite system (3.6) in the form

$$\begin{aligned} \frac{dx}{dt} &= \widetilde{A}x + \left(A(\varphi_t(\varphi)) - \widetilde{A}\right)x, \quad t \neq t_i(\varphi), \\ \Delta x|_{t=t_i(\varphi)} &= \widetilde{B}x + \left(B(\varphi_{t_i(\varphi)}(\varphi)) - \widetilde{B}\right)x. \end{aligned} \tag{3.8}$$

The fundamental matrix $\Omega_\tau^t(\varphi)$ of the system (3.6) may be represented in the following way [4]:

$$\begin{aligned} \Omega_\tau^t(\varphi) = X_\tau^t(\varphi) &+ \int_\tau^t X_s^t(\varphi)\left(A(\varphi_s(\varphi)) - \widetilde{A}\right)\Omega_s^t(\varphi)\,ds \\ &+ \sum_{\tau \le t_i(\varphi) < t} X_{t_i(\varphi)}^t(\varphi)\left(B(\varphi_{t_i(\varphi)}(\varphi)) - \widetilde{B}\right)\Omega_\tau^{t_i(\varphi)}(\varphi), \end{aligned} \tag{3.9}$$

to Contents, page 1

where $X_\tau^t(\varphi)$ is the fundamental matrix of the homogeneous impulsive system with constant coefficients

$$\begin{aligned} \frac{dx}{dt} &= \tilde{A}x, \quad t \neq t_i(\varphi), \\ \Delta x|_{t=t_i(\varphi)} &= \tilde{B}x \end{aligned} \tag{3.10}$$

that depends on $\varphi \in \mathcal{T}^m$ as a parameter. Taking into account that the matrix $X_\tau^t(\varphi)$ satisfies the estimate [14]

$$\|X_\tau^t(\varphi)\| \le Ke^{-\mu(t-\tau)}, \quad t \ge \tau \tag{3.11}$$

for every $\varphi \in \mathcal{T}^m$ and some $K \ge 1$, where $\gamma + p\ln\alpha < -\mu < 0$, we obtain

$$\begin{aligned} \|\Omega_\tau^t(\varphi)\| \le Ke^{-\mu(t-\tau)} &+ \int_\tau^t Ke^{-\mu(t-s)}\left\|A(\varphi_s(\varphi)) - \tilde{A}\right\|\|\Omega_\tau^s(\varphi)\|ds \\ &+ \sum_{\tau \le t_i(\varphi) < t} Ke^{-\mu(t-t_i(\varphi))}\left\|B(\varphi_{t_i(\varphi)}(\varphi)) - \tilde{B}\right\|\left\|\Omega_\tau^{t_i(\varphi)}(\varphi)\right\|. \end{aligned} \tag{3.12}$$

It follows from (3.2) that for arbitrary small ε_A and ε_B, there exists a moment T such that

$$\begin{aligned} \left\|A(\varphi_t(\varphi)) - \tilde{A}\right\| &\le \varepsilon_A, \\ \left\|B(\varphi_t(\varphi)) - \tilde{B}\right\| &\le \varepsilon_B \end{aligned} \tag{3.13}$$

for all $t \ge T$. Hence, multiplying (3.12) by $e^{\mu(t-\tau)}$, utilizing (3.13), and weakening the inequality, we obtain

$$\begin{aligned} e^{\mu(t-\tau)}\|\Omega_\tau^t(\varphi)\| \le K &+ \int_\tau^T Ke^{\mu(s-\tau)}\left\|A(\varphi_s(\varphi)) - \tilde{A}\right\|\|\Omega_\tau^s(\varphi)\|ds \\ &+ \int_\tau^t K\varepsilon_A e^{\mu(s-\tau)}\|\Omega_\tau^s(\varphi)\|ds \\ &+ \sum_{\tau \le t_i(\varphi) < T} Ke^{\mu(t_i(\varphi)-\tau)}\left\|B(\varphi_{t_i(\varphi)}(\varphi)) - \tilde{B}\right\|\left\|\Omega_\tau^{t_i(\varphi)}(\varphi)\right\| \\ &+ \sum_{\tau \le t_i(\varphi) < t} K\varepsilon_B e^{\mu(t_i(\varphi)-\tau)}\left\|\Omega_\tau^{t_i(\varphi)}(\varphi)\right\|. \end{aligned} \tag{3.14}$$

Using the Gronwall-Bellman inequality for piecewise continuous functions [4], we obtain the estimate for the fundamental matrix $\Omega_\tau^t(\varphi)$ of the system (3.6)

$$\|\Omega_\tau^t(\varphi)\| \le K_1 e^{-(\mu - K\varepsilon_A - p\ln(1+K\varepsilon_B))(t-\tau)}, \tag{3.15}$$

to Contents, page 1

where

$$K_1 = K + \int_{\tau}^{T} Ke^{\mu(s-\tau)} \left\| A(\varphi_s(\varphi)) - \widetilde{A} \right\| \left\| \Omega_{\tau}^{s}(\varphi) \right\| ds + \sum_{\tau \le t_i(\varphi) < T} Ke^{\mu(t_i(\varphi)-\tau)} \left\| B(\varphi_{t_i(\varphi)}(\varphi)) - \widetilde{B} \right\| \left\| \Omega_{\tau}^{t_i(\varphi)}(\varphi) \right\|. \tag{3.16}$$

Choosing ε_A and ε_B so that $\mu > K\varepsilon_A + p\ln(1 + K\varepsilon_B)$, the following estimate holds

$$\left\| \Omega_{\tau}^{t}(\varphi) \right\| \le K_1 e^{-\gamma_1(t-\tau)} \tag{3.17}$$

for all $t \ge \tau$ and some $K_1 \ge 1$, $\gamma_1 > 0$.

Estimate (3.17) is a sufficient condition for the existence and asymptotic stability of an invariant set of system (2.1). Indeed, it is easy to verify that invariant set $x = u(\varphi)$ of the system (2.1) may be represented as

$$u(\varphi) = \int_{-\infty}^{0} \Omega_{\tau}^{0}(\varphi) f(\varphi_{\tau}(\varphi)) d\tau + \sum_{t_i(\varphi) < 0} \Omega_{t_i(\varphi)}^{0}(\varphi) g(\varphi_{t_i(\varphi)}(\varphi)). \tag{3.18}$$

The integral and the sum from (3.18) converge since inequality (3.17) holds and limit (3.4) exists. Utilizing the properties (3.7) of the matrix $\Omega_{\tau}^{t}(\varphi)$ (2.4), and (2.6), one can show that the function $u(\varphi_t(\varphi))$ satisfies the equation

$$\frac{dx}{dt} = A(\varphi_t(\varphi))x + f(\varphi_t(\varphi)) \tag{3.19}$$

for $t \ne t_i(\varphi)$ and has discontinuities $B(\varphi_{t_i(\varphi)}(\varphi))u(\varphi_{t_i(\varphi)}(\varphi)) + g(\varphi_{t_i(\varphi)}(\varphi))$ at the points $t = t_i(\varphi)$. It means that the function $x(t,\varphi) = u(\varphi_t(\varphi))$ is a solution of the system (2.7). Hence, $u(\varphi)$ defines the invariant set of system (2.1).

Let us prove the asymptotic stability of the invariant set. Let $x = x(t,\varphi)$ be an arbitrary solutions of the system (2.7), and $x^* = u(\varphi_t(\varphi))$ is the solution that belongs to the invariant set. The difference of these solutions admits the representation

$$x(t,\varphi) - u(\varphi_t(\varphi)) = \Omega_0^t(\varphi)(x(0,\varphi) - u(\varphi)). \tag{3.20}$$

Taking into account estimate (3.17), the following limit exists

$$\lim_{t\to\infty} \left\| x(t,\varphi) - u(\varphi_t(\varphi)) \right\| = 0. \tag{3.21}$$

It proves the asymptotic stability of the invariant set $x = u(\varphi)$. □

to Contents, page 1

4. Perturbation Theory

Let us show that small perturbations of the right-hand side of the system (2.1) do not ruin the invariant set. Let $A_1(\varphi)$ and $B_1(\varphi)$ be continuous 2π-periodic with respect to each of the components φ_v, $v = 1, \ldots, m$ square matrices. Consider the perturbed system

$$\begin{gathered}\frac{d\varphi}{dt} = a(\varphi),\\ \frac{dx}{dt} = (A(\varphi) + A_1(\varphi))x + f(\varphi), \quad \varphi \notin \Gamma,\\ \Delta x|_{\varphi\in\Gamma} = (B(\varphi) + B_1(\varphi))x + g(\varphi).\end{gathered} \tag{4.1}$$

Theorem 4.1. *Let the moments of impulsive perturbations $\{t_i(\varphi)\}$ be such that uniformly with respect to $t \in \mathbb{R}$, there exists a finite limit*

$$\lim_{\tilde{T}\to\infty} \frac{i\left(t, t + \tilde{T}\right)}{\tilde{T}} = p \tag{4.2}$$

and the following inequality holds

$$\gamma + p \ln \alpha < 0. \tag{4.3}$$

Then there exist sufficiently small constants $a_1 > 0$ and $b_1 > 0$ such that for any continuous 2π-periodic with respect to each of the components φ_v, $v = 1, \ldots, m$ functions $A_1(\varphi)$ and $B_1(\varphi)$ such that

$$\begin{gathered}\max_{\varphi\in\mathcal{T}^m} \|A_1(\varphi)\| \le a_1,\\ \max_{\varphi\in\mathcal{T}^m} \|B_1(\varphi)\| \le b_1,\end{gathered} \tag{4.4}$$

system (4.1) *has an asymptotically stable invariant set.*

Proof. The constants a_1 and b_1 exist since the matrices $A_1(\varphi)$ and $B_1(\varphi)$ are continuous functions defined in the torus $\mathcal{T}^m$, which is a compact manifold.

Consider the impulsive system that corresponds to system(4.1)

$$\begin{gathered}\frac{dx}{dt} = A(\varphi_t(\varphi))x + A_1(\varphi_t(\varphi))x, \quad t \ne t_i(\varphi),\\ \Delta x|_{t=t_i(\varphi)} = B(\varphi_{t_i(\varphi)}(\varphi))x + B_1(\varphi_{t_i(\varphi)}(\varphi))x\end{gathered} \tag{4.5}$$

to Contents, page 1

that depends on $\varphi \in \mathcal{T}^m$ as a parameter. The fundamental matrix $\Psi_\tau^t(\varphi)$ of the system (4.5) may be represented in the following way

$$\Psi_\tau^t(\varphi) = \Omega_\tau^t(\varphi) + \int_\tau^t \Omega_s^t(\varphi) A_1(\varphi_s(\varphi)) \Psi_s^t(\varphi) ds + \sum_{\tau \le t_i(\varphi) < t} \Omega_{t_i(\varphi)}^t(\varphi) B_1(\varphi_{t_i(\varphi)}(\varphi)) \Psi_\tau^{t_i(\varphi)}(\varphi), \tag{4.6}$$

where $\Omega_\tau^t(\varphi)$ is the fundamental matrix of the system (3.6). Then taking estimate (3.17) into account,

$$e^{\gamma_1(t-\tau)} \left\| \Psi_\tau^t(\varphi) \right\| \le K_1 + \int_\tau^t K_1 a_1 e^{\gamma_1(s-\tau)} \left\| \Psi_\tau^s(\varphi) \right\| ds + \sum_{\tau \le t_i(\varphi) < t} K_1 b_1 e^{\gamma_1(t_i(\varphi)-\tau)} \left\| \Psi_\tau^{t_i(\varphi)}(\varphi) \right\|. \tag{4.7}$$

Using the Gronwall-Bellman inequality for piecewise continuous functions, we obtain the estimate for the fundamental matrix $\Psi_\tau^t(\varphi)$ of the system (4.5)

$$\left\| \Psi_\tau^t(\varphi) \right\| \le K_1 e^{-(\gamma_1 - K_1 a_1 - p \ln(1+K_1 b_1))(t-\tau)}. \tag{4.8}$$

Let the constants a_1 and b_1 be such that $\gamma_1 > K_1 a_1 + p \ln(1 + K_1 b_1)$. Hence, the matrix $\Psi_\tau^t(\varphi)$ satisfies the estimate

$$\left\| \Psi_\tau^t(\varphi) \right\| \le K_2 e^{-\gamma_2(t-\tau)} \tag{4.9}$$

for all $t \ge \tau$ and some $K_2 \ge 1$, $\gamma_2 > 0$. As in Theorem 3.1, from estimate (4.9), we conclude that the system (4.1) has an asymptotically stable invariant set $x = u(\varphi)$, which admits the representation

$$u(\varphi) = \int_{-\infty}^0 \Psi_\tau^0(\varphi) f(\varphi_\tau(\varphi)) d\tau + \sum_{t_i(\varphi) < 0} \Psi_{t_i(\varphi)}^0(\varphi) g(\varphi_{t_i(\varphi)}(\varphi)). \tag{4.10}$$

□

Consider the nonlinear system of differential equations with impulsive perturbations of the form

$$\begin{aligned} \frac{d\varphi}{dt} &= a(\varphi), \\ \frac{dx}{dt} &= F(\varphi, x), \quad \varphi \notin \Gamma, \\ \Delta x\big|_{\varphi \in \Gamma} &= I(\varphi, x), \end{aligned} \tag{4.11}$$

to Contents, page 1

where $\varphi \in \mathcal{T}^m$, $x \in \overline{J}_h$, $a(\varphi)$ is a continuous 2π-periodic with respect to each of the components φ_v, $v = 1,\dots,m$ vector function and satisfies Lipschitz conditions (2.2); $F(\varphi,x)$ and $I(\varphi,x)$ are continuous 2π-periodic with respect to each of the components φ_v, $v = 1,\dots,m$ functions that have continuous partial derivatives with respect to x up to the second order inclusively. Taking these assumptions into account, system (4.11) may be rewritten in the following form:

$$\begin{aligned} \frac{d\varphi}{dt} &= a(\varphi), \\ \frac{dx}{dt} &= A_0(\varphi)x + A_1(\varphi,x)x + f(\varphi), \quad \varphi \notin \Gamma, \\ \Delta x|_{\varphi\in\Gamma} &= B_0(\varphi)x + B_1(\varphi,x)x + g(\varphi), \end{aligned} \tag{4.12}$$

where

$$A(\varphi,x) = \int_0^1 \frac{\partial F(\varphi,\tau x)}{\partial(\tau x)}d\tau, \qquad B(\varphi,x) = \int_0^1 \frac{\partial I(\varphi,\tau x)}{\partial(\tau x)}d\tau, \tag{4.13}$$

$A_0(\varphi) = A(\varphi,0)$, $A_1(\varphi,x) = A(\varphi,x) - A(\varphi,0)$, $B_0(\varphi) = B(\varphi,0)$, $B_1(\varphi,x) = B(\varphi,x) - B(\varphi,0)$, $f(\varphi) = F(\varphi,0)$, and $g(\varphi) = I(\varphi,0)$. We assume that the matrices $A_0(\varphi)$ and $B_0(\varphi)$ are constant in the domain Ω:

$$\begin{aligned} A_0(\varphi)\big|_{\varphi\in\Omega} &= \widetilde{A}, \\ B_0(\varphi)\big|_{\varphi\in\Omega} &= \widetilde{B} \end{aligned} \tag{4.14}$$

and the inequality $\gamma + p\ln\alpha < 0$ holds.

We will construct the invariant set of system (4.12) using an iteration method proposed in [8]. As initial invariant set M_0, we consider the set $x = 0$, as M_k—the invariant set of the system

$$\begin{aligned} \frac{d\varphi}{dt} &= a(\varphi), \\ \frac{dx}{dt} &= A_0(\varphi)x + A_1(\varphi,u_{k-1}(\varphi))x + f(\varphi), \quad \varphi \notin \Gamma, \\ \Delta x|_{\varphi\in\Gamma} &= B_0(\varphi)x + B_1(\varphi,u_{k-1}(\varphi))x + g(\varphi), \end{aligned} \tag{4.15}$$

where $x = u_{k-1}(\varphi)$ is the invariant set on $(k-1)$-step.

Using Theorem 4.1, the invariant set $x = u_k(\varphi)$, $k = 1,2,\dots$ may be represented as

$$u_k(\varphi) = \int_{-\infty}^0 \Psi_\tau^0(\varphi,k) f(\varphi_\tau(\varphi))d\tau + \sum_{t_i(\varphi)<0} \Psi^0_{t_i(\varphi)+0}(\varphi,k) g(\varphi_{t_i(\varphi)}(\varphi)), \tag{4.16}$$

to Contents, page 1

where $\Psi_\tau^t(\varphi,k)$ is the fundamental matrix of the homogeneous system

$$\begin{aligned}\frac{dx}{dt} &= \left(A_0(\varphi_t(\varphi)) + A_1(\varphi_t(\varphi), u_{k-1}(\varphi_t(\varphi)))\right)x, \quad t \neq t_i(\varphi),\\ \Delta x|_{t=t_i(\varphi)} &= \left(B_0(\varphi_t(\varphi)) + B_1(\varphi_t(\varphi), u_{k-1}(\varphi_t(\varphi)))\right)x\end{aligned} \tag{4.17}$$

that depends on $\varphi \in \mathcal{T}^m$ as a parameter and satisfies the estimate

$$\left\|\Psi_\tau^t(\varphi,k)\right\| \leq K_2 e^{-\gamma_2(t-\tau)} \tag{4.18}$$

for all $t \geq \tau$ and some $K_2 \geq 1$, $\gamma_2 > 0$ only if

$$\begin{aligned}&\max_{\varphi\in\mathcal{T}^m}\left\|A_1(\varphi, u_{k-1}(\varphi))\right\| \leq a_1,\\ &\max_{\varphi\in\mathcal{T}^m}\left\|B_1(\varphi, u_{k-1}(\varphi))\right\| \leq b_1.\end{aligned} \tag{4.19}$$

Let us prove that the invariant sets $x = u_k(\varphi)$ belong to the domain $\overline{J}_h$. Denote by

$$\begin{aligned}&\max_{\varphi\in\mathcal{T}^m}\|f(\varphi)\| \leq M_f,\\ &\max_{\varphi\in\mathcal{T}^m}\|g(\varphi)\| \leq M_g.\end{aligned} \tag{4.20}$$

Since the torus $\mathcal{T}^m$ is a compact manifold, such constants M_f and M_g exist. Analogously to [4], using the representation (4.16) and estimate (4.18), we obtain that

$$\|u_k(\varphi)\| \leq \frac{K_2}{\gamma_2}M_f + \frac{K_2}{1-e^{-\gamma_2\theta_1}}M_g, \tag{4.21}$$

where θ_1 is a minimum gap between moments of impulsive actions. Condition (3.4) guarantees that such constant θ_1 exists. Assume that the constants K_2 and γ_2 are such that $\|u(\varphi)\| \leq h$.

Let us obtain the conditions for the convergence of the sequence $\{u_k(\varphi)\}$. For this purpose, we estimate the difference $w_{k+1}(\varphi) = u_{k+1}(\varphi) - u_k(\varphi)$ and take into account that the functions $u_k(\varphi_t(\varphi))$ satisfy the relations

$$\begin{aligned}\frac{d}{dt}u_k(\varphi_t(\varphi)) &= \left(A_0(\varphi_t(\varphi)) + A_1(\varphi_t(\varphi), u_{k-1}(\varphi_t(\varphi)))\right)\\ &\quad\times u_k(\varphi_t(\varphi)) + f(\varphi_t(\varphi)), \quad t \neq t_i(\varphi),\\ \Delta u_k(\varphi_t(\varphi))|_{t=t_i(\varphi)} &= \left(B_0(\varphi_t(\varphi)) + B_1(\varphi_t(\varphi), u_{k-1}(\varphi_t(\varphi)))\right)\\ &\quad\times u_k(\varphi_t(\varphi)) + g(\varphi_t(\varphi))\end{aligned} \tag{4.22}$$

to Contents, page 1

for all $\varphi \in \mathcal{T}^m$, $k = 1, 2, \ldots$. Hence, the difference $w_{k+1}(\varphi) = u_{k+1}(\varphi) - u_k(\varphi)$ is the invariant set of the linear impulsive system

$$\frac{d\varphi}{dt} = a(\varphi),$$
$$\frac{dx}{dt} = (A_0(\varphi) + A_1(\varphi, u_k(\varphi)))x + (A_1(\varphi, u_k(\varphi)) - A_1(\varphi, u_{k-1}(\varphi)))u_k(\varphi), \quad \varphi \notin \Gamma,$$
$$\Delta x|_{\varphi\in\Gamma} = (B_0(\varphi) + B_1(\varphi, u_k(\varphi)))x + (B_1(\varphi, u_k(\varphi)) - B_1(\varphi, u_{k-1}(\varphi)))u_k(\varphi). \tag{4.23}$$

Then, taking (4.21) into account,

$$\begin{aligned}\max_{\varphi\in\mathcal{T}^m}\|u_{k+1}(\varphi) - u_k(\varphi)\| \le{}& \frac{K_2}{\gamma_2}\|A_1(\varphi, u_k(\varphi)) - A_1(\varphi, u_{k-1}(\varphi))\|\|u_k(\varphi)\| \\ &+ \frac{K_2}{1 - e^{-\gamma_2\theta}}\|B_1(\varphi, u_k(\varphi)) - B_1(\varphi, u_{k-1}(\varphi))\|\|u_k(\varphi)\|.\end{aligned} \tag{4.24}$$

Let the functions $A_1(\varphi, x)$ and $B_1(\varphi, x)$ satisfy the Lipschitz condition with constants L_A and L_B correspondingly. Then

$$\begin{aligned}\max_{\varphi\in\mathcal{T}^m}\|u_{k+1}(\varphi) - u_k(\varphi)\| &\le \frac{K_2}{\gamma_2}L_A h\|u_k(\varphi) - u_{k-1}(\varphi)\| + \frac{K_2}{1 - e^{-\gamma_2\theta}}L_B h\|u_k(\varphi) - u_{k-1}(\varphi)\| \\ &= \left(\frac{K_2 h}{\gamma_2}L_A + \frac{K_2 h}{1 - e^{-\gamma_2\theta}}L_B\right)\|u_k(\varphi) - u_{k-1}(\varphi)\|.\end{aligned} \tag{4.25}$$

Assuming that the constants L_A and L_B are so small that

$$\frac{K_2 h}{\gamma_2}L_A + \frac{K_2 h}{1 - e^{-\gamma_2\theta}}L_B < 1, \tag{4.26}$$

we conclude that the sequence $\{u_k(\varphi)\}$ converges uniformly with respect to $\varphi \in \mathcal{T}^m$ and

$$\lim_{k\to\infty} u_k(\varphi) = u(\varphi). \tag{4.27}$$

Thus, the invariant set $x = u(\varphi)$ admits the representation

$$u(\varphi) = \int_{-\infty}^{0} \Psi_\tau^0(\varphi) f(\varphi_\tau(\varphi))d\tau + \sum_{t_i(\varphi)<0} \Psi_{t_i(\varphi)}^0(\varphi) g(\varphi_{t_i(\varphi)}(\varphi)), \tag{4.28}$$

to Contents, page 1

where $\Psi_\tau^t(\varphi)$ is the fundamental matrix of the homogeneous system

$$\begin{aligned} \frac{dx}{dt} &= (A(\varphi_t(\varphi)) + A_1(\varphi_t(\varphi), u(\varphi_t(\varphi))))x, \quad t \neq t_i(\varphi), \\ \Delta x|_{t=t_i(\varphi)} &= (B(\varphi_t(\varphi)) + B(\varphi_t(\varphi), u(\varphi_t(\varphi))))x \end{aligned} \tag{4.29}$$

that depends on $\varphi \in \mathcal{T}^m$ as a parameter and satisfies the estimation

$$\left\|\Psi_\tau^t(\varphi)\right\| \leq K_2 e^{-\gamma_2(t-\tau)} \tag{4.30}$$

for all $t \geq \tau$ and some $K_2 \geq 1$, $\gamma_2 > 0$. The following assertion has been proved.

Theorem 4.2. *Let the matrices $A_0(\varphi)$ and $B_0(\varphi)$ be constant in the domain Ω:*

$$\begin{aligned} A_0(\varphi)\big|_{\varphi\in\Omega} &= \widetilde{A}, \\ B_0(\varphi)\big|_{\varphi\in\Omega} &= \widetilde{B}, \end{aligned} \tag{4.31}$$

uniformly with respect to $t \in \mathbb{R}$, there exists a finite limit

$$\lim_{\widetilde{T}\to\infty} \frac{i\left(t, t+\widetilde{T}\right)}{\widetilde{T}} = p \tag{4.32}$$

and the following inequality holds

$$\gamma + p \ln \alpha < 0, \tag{4.33}$$

where

$$\begin{aligned} \gamma &= \max_{j=1,\ldots,n} \operatorname{Re} \lambda_j\left(\widetilde{A}\right), \\ \alpha^2 &= \max_{j=1,\ldots,n} \lambda_j\left(\left(E + \widetilde{B}\right)^T \left(E + \widetilde{B}\right)\right). \end{aligned} \tag{4.34}$$

Then there exist sufficiently small constants a_1 and b_1 and sufficiently small Lipschitz constants L_A and L_B such that for any continuous 2π-periodic with respect to each of the components φ_v, $v = 1, \ldots, m$ matrices $F(\varphi, x)$ and $I(\varphi, x)$, which have continuous partial derivatives with respect to x up to the second order inclusively, such that

$$\begin{aligned} \max_{\varphi\in\mathcal{T}^m, x\in\overline{J}_h} \left\|A_1(\varphi, x)\right\| &\leq a_1, \\ \max_{\varphi\in\mathcal{T}^m, x\in\overline{J}_h} \left\|B_1(\varphi, x)\right\| &\leq b_1 \end{aligned} \tag{4.35}$$

to Contents, page 1

and for any $x', x'' \in \overline{J}_h$

$$\begin{aligned} \|A_1(\varphi, x') - A_1(\varphi, x'')\| &\le L_A \|x' - x''\|, \\ \|B_1(\varphi, x') - B_1(\varphi, x'')\| &\le L_B \|x' - x''\|, \end{aligned} \tag{4.36}$$

system (4.11) *has an asymptotically stable invariant set.*

5. Conclusion

In summary, we have obtained sufficient conditions for the existence and asymptotic stability of invariant sets of a linear impulsive system of differential equations defined in $\mathcal{T}^m \times E^n$ that has specific properties in the ω-limit set Ω of the trajectories $\varphi_t(\varphi)$. We have proved that it is sufficient to impose some restrictions on system (2.1) only in the domain Ω to guarantee the existence and asymptotic stability of the invariant set.

References

[1] M. Akhmet, *Principles of Discontinuous Dynamical Systems*, Springer, New York, NY, USA, 2010.

[2] M. Benchohra, J. Henderson, and S. Ntouyas, *Impulsive Differential Equations and Inclusions*, vol. 2 of *Contemporary Mathematics and Its Applications*, Hindawi Publishing Corporation, New York, NY, USA, 2006.

[3] V. Lakshmikantham, D. D. Baĭnov, and P. S. Simeonov, *Theory of Impulsive Differential Equations*, vol. 6 of *Series in Modern Applied Mathematics*, World Scientific, Teaneck, NJ, USA, 1989.

[4] A. M. Samoĭlenko and N. A. Perestyuk, *Impulsive Differential Equations*, vol. 14 of *World Scientific Series on Nonlinear Science. Series A: Monographs and Treatises*, World Scientific, River Edge, NJ, USA, 1995.

[5] J. Moser, "On the theory of quasiperiodic motions," *SIAM Review*, vol. 8, no. 2, pp. 145–172, 1966.

[6] N. Kryloff and N. Bogoliubov, *Introduction to Non-Linear Mechanics*, Annals of Mathematics Studies, no. 11, Princeton University Press, Princeton, NJ, USA, 1943.

[7] N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Non-Linear Oscillations*, International Monographs on Advanced Mathematics and Physics, Gordon and Breach Science, New York, NY, USA, 1961.

[8] A. M. Samoĭlenko, *Elements of the Mathematical Theory of Multi-Frequency Oscillations*, vol. 71 of *Mathematics and Its Applications (Soviet Series)*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.

[9] S. I. Dudzyaniĭ and M. O. Perestyuk, "On the stability of a trivial invariant torus of a class of systems with impulse perturbation," *Ukrainian Mathematical Journal*, vol. 50, no. 3, pp. 338–349, 1998.

[10] M. O. Perestyuk and P. V. Feketa, "Invariant manifolds of a class of systems of differential equations with impulse perturbation," *Nonlinear Oscillations*, vol. 13, no. 2, pp. 240–273, 2010.

[11] K. Schneider, S. I. Kostadinov, and G. T. Stamov, "Integral manifolds of impulsive differential equations defined on torus," *Proceedings of the Japan Academy, Series A*, vol. 75, no. 4, pp. 53–57, 1999.

[12] V. I. Tkachenko, "The Green function and conditions for the existence of invariant sets of sampled-data systems," *Ukrainian Mathematical Journal*, vol. 41, no. 10, pp. 1379–1383, 1989.

[13] M. O. Perestyuk and S. I. Baloga, "Existence of an invariant torus for a class of systems of differential equations," *Nonlinear Oscillations*, vol. 11, no. 4, pp. 520–529, 2008.

[14] A. M. Samoilenko and N. A. Perestyuk, "On stability of the solutions of systems with impulsive perturbations," *Differential Equations*, vol. 17, no. 11, pp. 1995–2001, 1981 (Russian).

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 178568, 19 pages
doi:10.1155/2011/178568

Research Article

New Stability Conditions for Linear Differential Equations with Several Delays

Leonid Berezansky[1] and Elena Braverman[2]

[1] *Department of Mathematics, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel*
[2] *Department of Mathematics and Statistics, University of Calgary, 2500 University Drive Northwest, Calgary, AB, Canada T2N 1N4*

Correspondence should be addressed to Elena Braverman, maelena@math.ucalgary.ca

Received 28 January 2011; Accepted 4 April 2011

Academic Editor: Josef Diblík

New explicit conditions of asymptotic and exponential stability are obtained for the scalar nonautonomous linear delay differential equation $\dot{x}(t) + \sum_{k=1}^{m} a_k(t)x(h_k(t)) = 0$ with measurable delays and coefficients. These results are compared to known stability tests.

1. Introduction

In this paper we continue the study of stability properties for the scalar linear differential equation with several delays and an arbitrary number of positive and negative coefficients

$$\dot{x}(t) + \sum_{k=1}^{m} a_k(t)x(h_k(t)) = 0, \quad t \geq t_0, \tag{1.1}$$

which was begun in [1–3]. Equation (1.1) and its special cases were intensively studied, for example, in [4–21]. In [2] we gave a review of stability tests obtained in these papers.

In almost all papers on stability of delay-differential equations coefficients and delays are assumed to be continuous, which is essentially used in the proofs of main results. In real-world problems, for example, in biological and ecological models with seasonal fluctuations of parameters and in economical models with investments, parameters of differential equations are not necessarily continuous.

There are also some mathematical reasons to consider differential equations without the assumption that parameters are continuous functions. One of the main methods to

to Contents, page 1

investigate impulsive differential equations is their reduction to a nonimpulsive differential equation with discontinuous coefficients. Similarly, difference equations can sometimes be reduced to the similar problems for delay-differential equations with discontinuous piecewise constant delays.

In paper [1] some problems for differential equations with several delays were reduced to similar problems for equations with one delay which generally is not continuous.

One of the purposes of this paper is to extend and partially improve most popular stability results for linear delay equations with continuous coefficients and delays to equations with measurable parameters.

Another purpose is to generalize some results of [1–3]. In these papers, the sum of coefficients was supposed to be separated from zero and delays were assumed to be bounded. So the results of these papers are not applicable, for example, to the following equations:

$$\begin{gathered} \dot{x}(t) + |\sin t| x(t-\tau) = 0, \\ \dot{x}(t) + \alpha(|\sin t| - \sin t) x(t-\tau) = 0, \\ \dot{x}(t) + \frac{1}{t} x(t) + \frac{\alpha}{t} x\left(\frac{t}{2}\right) = 0. \end{gathered} \tag{1.2}$$

In most results of the present paper these restrictions are omitted, so we can consider all the equations mentioned above. Besides, necessary stability conditions (probably for the first time) are obtained for (1.1) with nonnegative coefficients and bounded delays. In particular, if this equation is exponentially stable then the ordinary differential equation

$$\dot{x}(t) + \sum_{k=1}^{m} a_k(t) x(t) = 0 \tag{1.3}$$

is also exponentially stable.

2. Preliminaries

We consider the scalar linear equation with several delays (1.1) for $t \geq t_0$ with the initial conditions (for any $t_0 \geq 0$)

$$x(t) = \varphi(t), \quad t < t_0, \qquad x(t_0) = x_0, \tag{2.1}$$

and under the following assumptions:

(a1) $a_k(t)$ are Lebesgue measurable essentially bounded on $[0, \infty)$ functions;

(a2) $h_k(t)$ are Lebesgue measurable functions,

$$h_k(t) \leq t, \qquad \limsup_{t \to \infty} h_k(t) = \infty; \tag{2.2}$$

(a3) $\varphi : (-\infty, t_0) \to R$ is a Borel measurable bounded function.

We assume conditions (a1)–(a3) hold for all equations throughout the paper.

to Contents, page 1

Definition 2.1. A locally absolutely continuous for $t \geq t_0$ function $x : R \to R$ is called *a solution* of problem (1.1), (2.1) if it satisfies (1.1) for almost all $t \in [t_0, \infty)$ and the equalities (2.1) for $t \leq t_0$.

Below we present a solution representation formula for the nonhomogeneous equation with locally Lebesgue integrable right-hand side $f(t)$:

$$\dot{x}(t) + \sum_{k=1}^{m} a_k(t)x(h_k(t)) = f(t), \quad t \geq t_0. \tag{2.3}$$

Definition 2.2. A solution $X(t,s)$ of the problem

$$\begin{gathered}\dot{x}(t) + \sum_{k=1}^{m} a_k(t)x(h_k(t)) = 0, \quad t \geq s \geq 0, \\ x(t) = 0, \quad t < s, \qquad x(s) = 1,\end{gathered} \tag{2.4}$$

is called *the fundamental function* of (1.1).

Lemma 2.3 (see [22, 23]). *Suppose conditions (a1)–(a3) hold. Then the solution of* (2.3), (2.1) *has the following form*

$$x(t) = X(t,t_0)x_0 - \int_{t_0}^{t} X(t,s)\sum_{k=1}^{m} a_k(s)\varphi(h_k(s))ds + \int_{t_0}^{t} X(t,s)f(s)ds, \tag{2.5}$$

where $\varphi(t) = 0$, $t \geq t_0$.

Definition 2.4 (see [22]). Equation (1.1) is *stable* if for any initial point t_0 and number $\varepsilon > 0$ there exists $\delta > 0$ such that the inequality $\sup_{t<t_0}|\varphi(t)| + |x(t_0)| < \delta$ implies $|x(t)| < \varepsilon$, $t \geq t_0$, for the solution of problem (1.1), (2.1).

Equation (1.1) is *asymptotically stable* if it is stable and all solutions of (1.1)-(2.1) for any initial point t_0 tend to zero as $t \to \infty$.

In particular, (1.1) is asymptotically stable if the fundamental function is uniformly bounded: $|X(t,s)| \leq K$, $t \geq s \geq 0$ and all solutions tend to zero as $t \to \infty$.

We apply in this paper only these two conditions of asymptotic stability.

Definition 2.5. Equation (1.1) is *(uniformly) exponentially stable*, if there exist $M > 0$, $\mu > 0$ such that the solution of problem (1.1), (2.1) has the estimate

$$|x(t)| \leq Me^{-\mu(t-t_0)}\left(|x(t_0)| + \sup_{t<t_0}\left|\varphi(t)\right|\right), \quad t \geq t_0, \tag{2.6}$$

where M and μ do not depend on t_0.

to Contents, page 1

Definition 2.6. The fundamental function $X(t,s)$ of (1.1) *has an exponential estimation* if there exist $K>0$, $\lambda>0$ such that

$$|X(t,s)| \leq Ke^{-\lambda(t-s)}, \quad t \geq s \geq 0. \tag{2.7}$$

For the linear (1.1) with bounded delays the last two definitions are equivalent. For unbounded delays estimation (2.7) implies asymptotic stability of (1.1).

Under our assumptions the exponential stability does not depend on values of equation parameters on any finite interval.

Lemma 2.7 (see [24, 25]). *Suppose $a_k(t) \geq 0$. If*

$$\int_{\max\{h(t),t_0\}}^{t} \sum_{i=1}^{m} a_i(s)ds \leq \frac{1}{e}, \quad h(t) = \min_k\{h_k(t)\}, \quad t \geq t_0, \tag{2.8}$$

or there exists $\lambda>0$, such that

$$\lambda \geq \sum_{k=1}^{m} A_k e^{\lambda\sigma_k}, \tag{2.9}$$

where

$$0 \leq a_k(t) \leq A_k, \quad t-h_k(t) \leq \sigma_k, \quad t \geq t_0, \tag{2.10}$$

then $X(t,s)>0$, $t \geq s \geq t_0$, where $X(t,s)$ is the fundamental function of (1.1).

Lemma 2.8 (see [3]). *Suppose $a_k(t) \geq 0$,*

$$\liminf_{t\to\infty} \sum_{k=1}^{m} a_k(t) > 0, \tag{2.11}$$

$$\limsup_{t\to\infty}(t-h_k(t)) < \infty, \quad k=1,\ldots,m, \tag{2.12}$$

and there exists $r(t) \leq t$ such that for sufficiently large t

$$\int_{r(t)}^{t} \sum_{k=1}^{m} a_k(s)ds \leq \frac{1}{e}. \tag{2.13}$$

If

$$\limsup_{t\to\infty} \sum_{k=1}^{m} \frac{a_k(t)}{\sum_{i=1}^{m} a_i(t)} \left| \int_{h_k(t)}^{r(t)} \sum_{i=1}^{m} a_i(s)ds \right| < 1, \tag{2.14}$$

then (1.1) *is exponentially stable.*

to Contents, page 1

Lemma 2.9 (see [3]). *Suppose* (2.12) *holds and there exists a set of indices* $I \subset \{1, \ldots, m\}$, *such that* $a_k(t) \geq 0$, $k \in I$,

$$\liminf_{t\to\infty} \sum_{k\in I} a_k(t) > 0, \tag{2.15}$$

and the fundamental function of the equation

$$\dot{x}(t) + \sum_{k\in I} a_k(t) x(h_k(t)) = 0 \tag{2.16}$$

is eventually positive. If

$$\limsup_{t\to\infty} \frac{\sum_{k\notin I} |a_k(t)|}{\sum_{k\in I} a_k(t)} < 1, \tag{2.17}$$

then (1.1) *is exponentially stable.*

The following lemma was obtained in [26, Corollary 2], see also [27].

Lemma 2.10. *Suppose for* (1.1) *condition* (2.12) *holds and this equation is exponentially stable. If*

$$\int_0^\infty \sum_{k=1}^n |b_k(s)| ds < \infty, \qquad \limsup_{t\to\infty} \left(t - g_k(t)\right) < \infty, \quad g_k(t) \leq t, \tag{2.18}$$

then the equation

$$\dot{x}(t) + \sum_{k=1}^m a_k(t) x(h_k(t)) + \sum_{k=1}^n b_k(t) x\left(g_k(t)\right) = 0 \tag{2.19}$$

is exponentially stable.

The following elementary result will be used in the paper.

Lemma 2.11. *The ordinary differential equation*

$$\dot{x}(t) + a(t)x(t) = 0 \tag{2.20}$$

is exponentially stable if and only if there exists $R > 0$ *such that*

$$\liminf_{t\to\infty} \int_t^{t+R} a(s) ds > 0. \tag{2.21}$$

to Contents, page 1

The following example illustrates that a stronger than (2.21) sufficient condition

$$\liminf_{t,s\to\infty}\frac{1}{t-s}\int_s^t a(\tau)d\tau > 0 \tag{2.22}$$

is not necessary for the exponential stability of the ordinary differential equation (2.20).

Example 2.12. Consider the equation

$$\dot{x}(t)+a(t)x(t)=0, \quad \text{where } a(t)=\begin{cases}1, & t\in[2n,2n+1),\\ 0, & t\in[2n+1,2n+2),\end{cases} \quad n=0,1,2,\ldots \tag{2.23}$$

Then lim inf in (2.22) equals zero, but $|X(t,s)| < ee^{-0.5(t-s)}$, so the equation is exponentially stable. Moreover, if we consider lim inf in (2.22) under the condition $t-s\geq R$, then it is still zero for any $R\leq 1$.

3. Main Results

Lemma 3.1. *Suppose* $a_k(t)\geq 0$, (2.11), (2.12) *hold and*

$$\limsup_{t\to\infty}\sum_{k=1}^m\frac{a_k(t)}{\sum_{i=1}^m a_i(t)}\int_{h_k(t)}^t\sum_{i=1}^m a_i(s)ds < 1+\frac{1}{e}. \tag{3.1}$$

Then (1.1) *is exponentially stable.*

Proof. By (2.11) there exists function $r(t)\leq t$ such that for sufficiently large t

$$\int_{r(t)}^t\sum_{k=1}^m a_k(s)ds=\frac{1}{e}. \tag{3.2}$$

For this function condition (2.14) has the form

$$\begin{aligned}&\limsup_{t\to\infty}\sum_{k=1}^m\frac{a_k(t)}{\sum_{i=1}^m a_i(t)}\left|\int_{h_k(t)}^t\sum_{i=1}^m a_i(s)ds-\int_{r(t)}^t\sum_{i=1}^m a_i(s)ds\right|\\ &=\limsup_{t\to\infty}\sum_{k=1}^m\frac{a_k(t)}{\sum_{i=1}^m a_i(t)}\left|\int_{h_k(t)}^t\sum_{i=1}^m a_i(s)ds-\frac{1}{e}\right|<1.\end{aligned} \tag{3.3}$$

The latter inequality follows from (3.1). The reference to Lemma 2.8 completes the proof. ◻

to Contents, page 1

Corollary 3.2. *Suppose* $a_k(t) \geq 0$, (2.11), (2.12) *hold and*

$$\limsup_{t\to\infty} \int_{\min_k\{h_k(t)\}}^{t} \sum_{i=1}^{m} a_i(s)ds < 1 + \frac{1}{e}. \tag{3.4}$$

Then (1.1) *is exponentially stable.*

The following theorem contains stability conditions for equations with unbounded delays. We also omit condition (2.11) in Lemma 3.1.

We recall that $b(t) > 0$ in the space of Lebesgue measurable essentially bounded functions means $b(t) \geq 0$ and $b(t) \neq 0$ almost everywhere.

Theorem 3.3. *Suppose* $a_k(t) \geq 0$, *condition* (3.1) *holds,* $\sum_{k=1}^{m} a_k(t) > 0$ *and*

$$\int_0^{\infty} \sum_{k=1}^{m} a_k(t)dt = \infty, \qquad \limsup_{t\to\infty} \int_{h_k(t)}^{t} \sum_{i=1}^{m} a_i(s)ds < \infty. \tag{3.5}$$

Then (1.1) *is asymptotically stable.*

If in addition there exists $R > 0$ *such that*

$$\liminf_{t\to\infty} \int_t^{t+R} \sum_{k=1}^{m} a_k(\tau)d\tau > 0 \tag{3.6}$$

then the fundamental function of (1.1) *has an exponential estimation.*

If condition (2.12) *also holds then* (1.1) *is exponentially stable.*

Proof. Let $s = p(t) := \int_0^t \sum_{k=1}^{m} a_k(\tau)d\tau$, $y(s) = x(t)$, where $p(t)$ is a strictly increasing function. Then $x(h_k(t)) = y(l_k(s))$, $l_k(s) \leq s$, $l_k(s) = \int_0^{h_k(t)} \sum_{i=1}^{m} a_i(\tau)d\tau$ and (1.1) can be rewritten in the form

$$\dot{y}(s) + \sum_{k=1}^{m} b_k(s)y(l_k(s)) = 0, \tag{3.7}$$

where $b_k(s) = a_k(t)/\sum_{i=1}^{m} a_i(t)$, $s - l_k(s) = \int_{h_k(t)}^{t} \sum_{i=1}^{m} a_i(\tau)d\tau$. Since $\sum_{k=1}^{m} b_k(s) = 1$ and $\limsup_{s\to\infty}(s - l_k(s)) < \infty$, then Lemma 3.1 can be applied to (3.7). We have

$$\begin{aligned}
&\limsup_{s\to\infty} \sum_{k=1}^{m} \frac{b_k(s)}{\sum_{i=1}^{m} b_i(s)} \int_{l_k(s)}^{s} \sum_{i=1}^{m} b_i(\tau)d\tau \\
&\quad = \limsup_{s\to\infty} \sum_{k=1}^{m} b_k(s)(s - l_k(s)) \\
&\quad = \limsup_{t\to\infty} \sum_{k=1}^{m} \frac{a_k(t)}{\sum_{i=1}^{m} a_i(t)} \int_{h_k(t)}^{t} \sum_{i=1}^{m} a_i(s)ds < 1 + \frac{1}{e}.
\end{aligned} \tag{3.8}$$

to Contents, page 1

By Lemma 3.1, (3.7) is exponentially stable. Due to the first equality in (3.5) $t \to \infty$ implies $s \to \infty$. Hence $\lim_{t\to\infty} x(t) = \lim_{s\to\infty} y(s) = 0$.

Equation (3.7) is exponentially stable, thus the fundamental function $Y(u,v)$ of (3.7) has an exponential estimation

$$|Y(u,v)| \le K e^{-\lambda(u-v)}, \quad u \ge v \ge 0, \tag{3.9}$$

with $K > 0$, $\lambda > 0$. Since $X(t,s) = Y(\int_0^t \sum_{k=1}^m a_k(\tau)d\tau, \int_0^s \sum_{k=1}^m a_k(\tau)d\tau)$, where $X(t,s)$ is the fundamental function of (1.1), then (3.9) yields

$$|X(t,s)| \le K \exp\left\{-\lambda \int_s^t \sum_{k=1}^m a_k(\tau)d\tau\right\}. \tag{3.10}$$

Hence $|X(t,s)| \le K$, $t \ge s \ge 0$, which together with $\lim_{t\to\infty} x(t) = 0$ yields that (1.1) is asymptotically stable.

Suppose now that (3.6) holds. Without loss of generality we can assume that for some $R > 0$, $\alpha > 0$ we have

$$\int_t^{t+R} \sum_{k=1}^m a_k(\tau)d\tau \ge \alpha > 0, \quad t \ge s \ge 0. \tag{3.11}$$

Hence

$$\exp\left\{-\lambda \int_s^t \sum_{k=1}^m a_k(\tau)d\tau\right\} \le \exp\left\{\lambda R \sup_{t\ge 0} \sum_{k=1}^m a_k(t)\right\} e^{-\lambda\alpha(t-s)/R}. \tag{3.12}$$

Thus, condition (3.6) implies the exponential estimate for $X(t,s)$.

The last statement of the theorem is evident. □

Remark 3.4. The substitution $s = p(t) := \int_0^t \sum_{k=1}^m a_k(\tau)d\tau$, $y(s) = x(t)$ was first used in [28].

Note that in [10, Lemma 2] this idea was extended to a more general equation

$$\dot{x}(t) + \int_{t_0}^t x(s) d_s r(t,s) = 0. \tag{3.13}$$

The ideas of [10] allow to generalize the results of the present paper to equations with a distributed delay.

Corollary 3.5. *Suppose* $a_k(t) \ge 0$, $\sum_{k=1}^m a_k(t) \equiv \alpha > 0$, *condition* (2.12) *holds and*

$$\limsup_{t\to\infty} \sum_{k=1}^m a_k(t)(t - h_k(t)) < 1 + \frac{1}{e}. \tag{3.14}$$

Then (1.1) *is exponentially stable.*

to Contents, page 1

Corollary 3.6. *Suppose* $a_k(t) = \alpha_k p(t)$, $\alpha_k > 0$, $p(t) > 0$, $\int_0^\infty p(t)dt = \infty$ *and*

$$\limsup_{t\to\infty} \sum_{k=1}^{m} \alpha_k \int_{h_k(t)}^{t} p(s)ds < 1 + \frac{1}{e}. \tag{3.15}$$

Then (1.1) *is asymptotically stable.*

If in addition there exists $R > 0$ *such that*

$$\liminf_{t\to\infty} \int_{t}^{t+R} p(\tau)d\tau > 0, \tag{3.16}$$

then the fundamental function of (1.1) *has an exponential estimation.*

If also (2.12) *holds then* (1.1) *is exponentially stable.*

Remark 3.7. Let us note that similar results for (3.13) were obtained in [10], see Corollary 3.4 and remark after it, Theorem 4 and Corollaries 4.1 and 4.2 in [10], where an analogue of condition (3.16) was applied. This allows to extend the results of the present paper to equations with a distributed delay.

Corollary 3.8. *Suppose* $a(t) \geq 0$, $b(t) \geq 0$, $a(t) + b(t) > 0$,

$$\begin{gathered}\int_0^\infty (a(t) + b(t))dt = \infty, \qquad \limsup_{t\to\infty} \int_{h(t)}^{t} (a(s) + b(s))ds < \infty, \\ \limsup_{t\to\infty} \frac{b(t)}{a(t) + b(t)} \int_{h(t)}^{t} (a(s) + b(s))ds < 1 + \frac{1}{e}.\end{gathered} \tag{3.17}$$

Then the following equation is asymptotically stable

$$\dot{x}(t) + a(t)x(t) + b(t)x(h(t)) = 0. \tag{3.18}$$

If in addition there exists $R > 0$ *such that* $\liminf_{t\to\infty} \int_t^{t+R} (a(\tau) + b(\tau))d\tau > 0$ *then the fundamental function of* (3.18) *has an exponential estimation.*

If also $\limsup_{t\to\infty}(t - h(t)) < \infty$ *then* (3.18) *is exponentially stable.*

In the following theorem we will omit the condition $\sum_{k=1}^m a_k(t) > 0$ of Theorem 3.3.

Theorem 3.9. *Suppose* $a_k(t) \geq 0$, *condition* (3.4) *and the first inequality in* (3.5) *hold. Then* (1.1) *is asymptotically stable.*

If in addition (3.6) *holds then the fundamental function of* (1.1) *has an exponential estimation.*

If also (2.12) *holds then* (1.1) *is exponentially stable.*

Proof. For simplicity suppose that $m = 2$ and consider the equation

$$\dot{x}(t) + a(t)x(h(t)) + b(t)x\big(g(t)\big) = 0, \tag{3.19}$$

to Contents, page 1

where $a(t) \geq 0$, $b(t) \geq 0$, $\int_0^\infty (a(s)+b(s))ds = \infty$ and there exist $t_0 \geq 0$, $\varepsilon > 0$ such that

$$\int_{\min\{h(t),g(t)\}}^{t} (a(s)+b(s))ds < 1 + \frac{1}{e} - \varepsilon, \quad t \geq t_0. \tag{3.20}$$

Let us find $t_1 \geq t_0$ such that $e^{-h(t)} < \varepsilon/4$, $e^{-g(t)} < \varepsilon/4$, $t \geq t_1$, such t_1 exists due to (a2). Then $\int_{\min\{h(t),g(t)\}}^{t} e^{-s}ds < \varepsilon/2$, $t \geq t_1$. Rewrite (3.19) in the form

$$\dot{x}(t) + \left(a(t)+e^{-t}\right)x(h(t)) + b(t)x\left(g(t)\right) - e^{-t}x(h(t)) = 0, \tag{3.21}$$

where $a(t)+b(t)+e^{-t} > 0$. After the substitution $s = \int_{t_1}^{t}(a(\tau)+b(\tau)+e^{-\tau})d\tau$, $y(s) = x(t)$, (3.21) has the form

$$\dot{y}(s) + \frac{a(t)+e^{-t}}{a(t)+b(t)+e^{-t}}y(l(s)) + \frac{b(t)}{a(t)+b(t)+e^{-t}}y\left(p(s)\right) - \frac{e^{-t}}{a(t)+b(t)+e^{-t}}y(l(s)) = 0, \tag{3.22}$$

where similar to the proof of Theorem 3.3

$$s - l(s) = \int_{h(t)}^{t} \left(a(\tau)+b(\tau)+e^{-\tau}\right)d\tau, \qquad s - p(s) = \int_{g(t)}^{t} \left(a(\tau)+b(\tau)+e^{-\tau}\right)d\tau. \tag{3.23}$$

First we will show that by Corollary 3.2 the equation

$$\dot{y}(s) + \frac{a(t)+e^{-t}}{a(t)+b(t)+e^{-t}}y(l(s)) + \frac{b(t)}{a(t)+b(t)+e^{-t}}y\left(p(s)\right) = 0 \tag{3.24}$$

is exponentially stable. Since $(a(t)+e^{-t})/(a(t)+b(t)+e^{-t}) + b(t)/(a(t)+b(t)+e^{-t}) = 1$, then (2.11) holds. Condition (3.20) implies (2.12). So we have to check only condition (3.4) where the sum under the integral is equal to 1. By (3.20), (3.23) we have

$$\begin{aligned}\int_{\min\{l(s),p(s)\}}^{s} 1ds = s - \min\{l(s),p(s)\}, \qquad s - l(s) &= \int_{h(t)}^{t} \left(a(\tau)+b(\tau)+e^{-\tau}\right)d\tau \\ &= \int_{h(t)}^{t} (a(\tau)+b(\tau))d\tau + \int_{h(t)}^{t} e^{-\tau}d\tau < 1 + \frac{1}{e} - \varepsilon + \frac{\varepsilon}{2} = 1 + \frac{1}{e} - \frac{\varepsilon}{2}, \quad t \geq t_1.\end{aligned} \tag{3.25}$$

The same calculations give $s - p(s) < 1 + (1/e) - \varepsilon/2$, thus condition (3.4) holds.

Hence (3.24) is exponentially stable.

We return now to (3.22), $t \geq t_1$. We have $ds = (a(t)+b(t)+e^{-t})dt$, then

$$\int_{t_1}^{\infty} \frac{e^{-t}}{a(t)+b(t)+e^{-t}}ds = \int_{t_1}^{\infty} \frac{e^{-t}}{a(t)+b(t)+e^{-t}}\left(a(t)+b(t)+e^{-t}\right)dt < \infty. \tag{3.26}$$

to Contents, page 1

By Lemma 2.10, (3.22) is exponentially stable. Since $t \to \infty$ implies $s \to \infty$ then $\lim_{t\to\infty} x(t) = \lim_{s\to\infty} y(s) = 0$, which completes the proof of the first part of the theorem. The rest of the proof is similar to the proof of Theorem 3.3. □

Corollary 3.10. *Suppose $a(t) \geq 0$, $\int_0^\infty a(t)dt = \infty$ and*

$$\limsup_{t\to\infty} \int_{h(t)}^{t} a(s)ds < 1 + \frac{1}{e}. \tag{3.27}$$

Then the equation

$$\dot{x}(t) + a(t)x(h(t)) = 0 \tag{3.28}$$

is asymptotically stable. If in addition condition (2.21) *holds then the fundamental function of* (3.28) *has an exponential estimation. If also* $\limsup_{t\to\infty}(t - h(t)) < \infty$ *then* (3.28) *is exponentially stable.*

Now consider (1.1), where only some of coefficients are nonnegative.

Theorem 3.11. *Suppose there exists a set of indices $I \subset \{1, \ldots, m\}$ such that $a_k(t) \geq 0$, $k \in I$,*

$$\int_0^\infty \sum_{k\in I} a_k(t)dt = \infty, \quad \limsup_{t\to\infty} \int_{h_k(t)}^{t} \sum_{i\in I} a_i(s)ds < \infty, \quad k = 1, \ldots, m, \tag{3.29}$$

$$\sum_{k\notin I} |a_k(t)| = 0, \quad t \in E, \qquad \limsup_{t\to\infty, t\notin E} \frac{\sum_{k\notin I} |a_k(t)|}{\sum_{k\in I} a_k(t)} < 1, \quad \text{where } E = \left\{ t \geq 0, \sum_{k\in I} a_k(t) = 0 \right\}. \tag{3.30}$$

If the fundamental function $X_0(t,s)$ of (2.16) *is eventually positive then all solutions of* (1.1) *tend to zero as $t \to \infty$.*

If in addition there exists $R > 0$ such that

$$\liminf_{t\to\infty} \int_{t}^{t+R} \sum_{k\in I} a_k(\tau)d\tau > 0 \tag{3.31}$$

then the fundamental function of (1.1) *has an exponential estimation.*

If condition (2.12) *also holds then* (1.1) *is exponentially stable.*

Proof. Without loss of generality we can assume $X_0(t,s) > 0$, $t \geq s \geq 0$. Rewrite (1.1) in the form

$$\dot{x}(t) + \sum_{k\in I} a_k(t)x(h_k(t)) + \sum_{k\notin I} a_k(t)x(h_k(t)) = 0. \tag{3.32}$$

to Contents, page 1

Suppose first that $\sum_{k\in I} a_k(t) \not\equiv 0$. After the substitution $s = p(t) := \int_0^t \sum_{k\in I} a_k(\tau)d\tau$, $y(s) = x(t)$ we have $x(h_k(t)) = y(l_k(s))$, $l_k(s) \le s$, $l_k(s) = \int_0^{h_k(t)} \sum_{i\in I} a_i(\tau)d\tau$, $k = 1, \dots, m$, and (1.1) can be rewritten in the form

$$\dot{y}(s) + \sum_{k=1}^{m} b_k(s)y(l_k(s)) = 0, \tag{3.33}$$

where $b_k(s) = a_k(t)/\sum_{i\in I} a_i(t)$. Denote by $Y_0(u,v)$ the fundamental function of the equation

$$\dot{y}(s) + \sum_{k\in I} b_k(s)y(l_k(s)) = 0. \tag{3.34}$$

We have

$$\begin{gathered} X_0(t,s) = Y_0\left(\int_0^t \sum_{k\in I} a_k(\tau)d\tau, \int_0^s \sum_{k\in I} a_k(\tau)d\tau\right), \\ Y_0(u,v) = X_0\left(p^{-1}(u), p^{-1}(v)\right) > 0, \quad u \ge v \ge 0. \end{gathered} \tag{3.35}$$

Let us check that other conditions of Lemma 2.9 hold for (3.33). Since $\sum_{k\in I} b_k(s) = 1$ then condition (2.15) is satisfied. In addition,

$$\limsup_{s\to\infty, p^{-1}(s)\notin E} \frac{\sum_{k\notin I}|b_k(s)|}{\sum_{k\in I} b_k(s)} = \limsup_{t\to\infty, t\notin E} \frac{\sum_{k\notin I}|a_k(t)|}{\sum_{k\in I} a_k(t)} < 1. \tag{3.36}$$

By Lemma 2.9, (3.33) is exponentially stable. Hence for any solution $x(t)$ of (1.1) we have $\lim_{t\to\infty} x(t) = \lim_{s\to\infty} y(s) = 0$. The end of the proof is similar to the proof of Theorem 3.9. In particular, to remove the condition $\sum_{k\in I} a_k(t) \not\equiv 0$ we rewrite the equation by adding the term e^{-t} to one of $a_k(t), k \in I$. □

Remark 3.12. Explicit positiveness conditions for the fundamental function were presented in Lemma 2.7.

Corollary 3.13. *Suppose*

$$\begin{gathered} a(t) \ge 0, \quad \int_0^\infty a(t)dt = \infty, \quad \limsup_{t\to\infty} \int_{g_k(t)}^t a(s)ds < \infty, \\ \sum_{k=1}^n |b_k(t)| = 0, \quad t \in E, \qquad \limsup_{t\to\infty, t\notin E} \frac{\sum_{k=1}^n |b_k(t)|}{a(t)} < 1, \end{gathered} \tag{3.37}$$

where $E = \{t \ge 0, a(t) = 0\}$. Then the equation

$$\dot{x}(t) + a(t)x(t) + \sum_{k=1}^n b_k(t)x(g_k(t)) = 0 \tag{3.38}$$

to Contents, page 1

is asymptotically stable. If in addition (2.21) *holds then the fundamental function of* (3.38) *has an exponential estimation. If also* $\limsup_{t\to\infty}(t - g_k(t)) < \infty$ *then* (3.38) *is exponentially stable.*

Theorem 3.14. *Suppose* $\int_0^\infty \sum_{k=1}^m |a_k(s)|ds < \infty$. *Then all solutions of* (1.1) *are bounded and* (1.1) *is not asymptotically stable.*

Proof. For the fundamental function of (1.1) we have the following estimation

$$|X(t,s)| \le \exp\left\{\int_s^t \sum_{k=1}^m |a_k(\tau)|d\tau\right\}. \tag{3.39}$$

Then by solution representation formula (2.5) for any solution $x(t)$ of (1.1) we have

$$\begin{aligned} |x(t)| &\le \exp\left\{\int_{t_0}^t \sum_{k=1}^m |a_k(s)|ds\right\}|x(t_0)| + \int_{t_0}^t \exp\left\{\int_s^t \sum_{k=1}^m |a_k(\tau)|d\tau\right\}\sum_{k=1}^m |a_k(s)|\big|\varphi(h_k(s))\big|ds \\ &\le \exp\left\{\int_{t_0}^\infty \sum_{k=1}^m |a_k(s)|ds\right\}\left(|x(t_0)| + \int_{t_0}^\infty \sum_{k=1}^m |a_k(s)|ds\|\varphi\|\right), \end{aligned} \tag{3.40}$$

where $\|\varphi\| = \max_{t<0}|\varphi(t)|$. Then $x(t)$ is a bounded function.

Moreover, $|X(t,s)| \le A := \exp\{\int_0^\infty \sum_{k=1}^m |a_k(s)|ds\}$, $t \ge s \ge 0$. Let us choose $t_0 \ge 0$ such that $\int_{t_0}^\infty \sum_{k=1}^m |a_k(s)|ds < 1/(2A)$, then $X_t'(t,t_0) + \sum_{k=1}^m a_k(t)X(h_k(t),t_0) = 0$, $X(t_0,t_0) = 1$ implies $X(t,t_0) \ge 1 - \int_{t_0}^\infty \sum_{k=1}^m |a_k(s)|Ads > 1 - A(1/(2A)) = 1/2$, thus $X(t,t_0)$ does not tend to zero, so (1.1) is not asymptotically stable. □

Theorems 3.11 and 3.14 imply the following results.

Corollary 3.15. *Suppose* $a_k(t) \ge 0$, *there exists a set of indices* $I \subset \{1,\dots,m\}$ *such that condition* (3.30) *and the second condition in* (3.29) *hold. Then all solutions of* (1.1) *are bounded.*

Proof. If $\int_0^\infty \sum_{k\in I} |a_k(t)|dt = \infty$, then all solutions of (1.1) are bounded by Theorem 3.11. Let $\int_0^\infty \sum_{k\in I} |a_k(t)|dt < \infty$. By (3.30) we have $\int_0^\infty \sum_{k\notin I} |a_k(t)|dt \le \int_0^\infty \sum_{k\in I} |a_k(t)|dt < \infty$. Then $\int_0^\infty \sum_{k=1}^m |a_k(t)|dt < \infty$. By Theorem 3.14 all solutions of (1.1) are bounded. □

Theorem 3.16. *Suppose* $a_k(t) \ge 0$. *If* (1.1) *is asymptotically stable, then the ordinary differential equation*

$$\dot{x}(t) + \left(\sum_{k=1}^m a_k(t)\right)x(t) = 0 \tag{3.41}$$

is also asymptotically stable. If in addition (2.12) *holds and* (1.1) *is exponentially stable, then* (3.41) *is also exponentially stable.*

to Contents, page 1

Proof. The solution of (3.41), with the initial condition $x(t_0) = x_0$, can be presented as $x(t) = x_0 \exp\{-\int_{t_0}^{t} \sum_{k=1}^{m} a_k(s)ds\}$, so (3.41) is asymptotically stable, as far as

$$\int_0^{\infty} \sum_{k=1}^{m} a_k(s)ds = \infty \tag{3.42}$$

and is exponentially stable if (3.6) holds (see Lemma 2.11).

If (3.42) does not hold, then by Theorem 3.14, (1.1) is not asymptotically stable.

Further, let us demonstrate that exponential stability of (1.1) really implies (3.6).

Suppose for the fundamental function of (1.1) inequality (2.7) holds and condition (3.6) is not satisfied. Then there exists a sequence $\{t_n\}$, $t_n \to \infty$, such that

$$\int_{t_n}^{t_n+n} \sum_{k=1}^{m} a_k(\tau)d\tau < \frac{1}{n} < \frac{1}{e}, \quad n \geq 3. \tag{3.43}$$

By (2.12) there exists $n_0 > 3$ such that $t - h_k(t) \leq n_0$, $k = 1, \ldots, m$. Lemma 2.7 implies $X(t, s) > 0$, $t_n \leq s \leq t \leq t_n + n$, $n \geq n_0$. Similar to the proof of Theorem 3.14 and using the inequality $1 - x \geq e^{-x}$, $x > 0$, we obtain

$$X(t_n, t_n + n) \geq 1 - \int_{t_n}^{t_n+n} \sum_{k=1}^{m} a_k(\tau)d\tau \geq \exp\left\{-\int_{t_n}^{t_n+n} \sum_{k=1}^{m} a_k(\tau)d\tau\right\} > e^{-1/n}. \tag{3.44}$$

Inequality (2.7) implies $|X(t_n + n, t_n)| \leq Ke^{-\lambda n}$. Hence $Ke^{-\lambda n} \geq e^{-1/n}$, $n \geq n_0$, or $K > e^{\lambda n - 1/3}$ for any $n \geq n_0$. The contradiction proves the theorem. □

Theorems 3.11 and 3.16 imply the following statement.

Corollary 3.17. *Suppose* $a_k(t) \geq 0$ *and the fundamental function of* (1.1) *is positive. Then* (1.1) *is asymptotically stable if and only if the ordinary differential equation* (3.41) *is asymptotically stable.*

If in addition (2.12) *holds then* (1.1) *is exponentially stable if and only if* (3.41) *is exponentially stable.*

4. Discussion and Examples

In paper [2] we gave a review of known stability tests for the linear equation (1.1). In this part we will compare the new results obtained in this paper with known stability conditions.

First let us compare the results of the present paper with our papers [1–3]. In all these three papers we apply the same method based on Bohl-Perron-type theorems and comparison with known exponentially stable equations.

In [1–3] we considered exponential stability only. Here we also give explicit conditions for asymptotic stability. For this type of stability, we omit the requirement that the delays are bounded and the sum of the coefficients is separated from zero. We also present some new stability tests, based on the results obtained in [3].

Compare now the results of the paper with some other known results [5–7, 9, 10, 22]. First of all we replace the constant $3/2$ in most of these tests by the constant $1 + 1/e$. Evidently

to Contents, page 1

$1 + 1/e = 1.3678 \cdots < 3/2$, so we have a worse constant, but it is an open problem to obtain $(3/2)$-stability results for equations with measurable coefficients and delays.

Consider now (3.28) with a single delay. This equation is well studied beginning with the classical stability result by Myshkis [29]. We present here several statements which cover most of known stability tests for this equation.

Statement 1 (see [5]). Suppose $a(t) \geq 0$, $h(t) \leq t$ are continuous functions and

$$\limsup_{t \to \infty} \int_{h(t)}^{t} a(s)ds \leq \frac{3}{2}. \tag{4.1}$$

Then all solutions of (3.28) are bounded.

If in addition

$$\liminf_{t \to \infty} \int_{h(t)}^{t} a(s)ds > 0, \tag{4.2}$$

and the strict inequality in (4.1) holds then (3.28) is exponentially stable.

Statement 2 (see [7]). Suppose $a(t) \geq 0$, $h(t) \leq t$ are continuous functions, the strict inequality (4.1) holds and $\int_0^\infty a(s)ds = \infty$. Then all solutions of (3.28) tend to zero as $t \to \infty$.

Statement 3 (see [9, 10]). Suppose $a(t) \geq 0$, $h(t) \leq t$ are measurable functions, $\int_0^\infty a(s)ds = \infty$, $A(t) = \int_0^t a(s)ds$ is a strictly monotone increasing function and

$$\limsup_{t \to \infty} \int_{h(t)}^{t} a(s)ds < \sup_{0<\omega<\pi/2} \left(\omega + \frac{1}{\Phi(\omega)}\right) \approx 1.45\ldots, \tag{4.3}$$

$\Phi(\omega) = \int_0^\infty u(t,\omega)dt$, where $u(t,\omega)$ is a solution of the initial value problem

$$\dot{y}(t) + y(t-\omega) = 0, \quad y(t) = 0, \quad t < 0, \quad y(0) = 1. \tag{4.4}$$

Then (3.28) is asymptotically stable.

Note that instead of the equation $\dot{y}(t) + y(t-\omega) = 0$ with a constant delay, the equation

$$\dot{y}(t) + y(t - \tau(t)) = 0 \tag{4.5}$$

can be used as the model equation. For example, the following results are valid.

Statement 4 (see [10]). Equation (4.5) is exponentially stable if $|\tau(t) - \omega| \leq k/\chi(\omega)$, where $k \in [0, \omega)$, $0 \leq \omega < \pi/2$ and $\chi(\omega) = \int_0^\infty |u(t,\omega)|dt$.

Obviously in this statement the delay can exceed 2.

Statement 5 (see [10]). Let $\tau(t) \leq k + \omega\{t/\omega\}$, where $k \in (0,1)$, $0 < \omega < 1$, $\{q\}$ is the fractional part of q. Then (4.5) is exponentially stable.

Here the delay $\tau(t)$ can be in the neighbourhood of ω which is close to 1.

Example 4.1. Consider the equation

$$\dot{x}(t) + \alpha(|\sin t| - \sin t)x(h(t)) = 0, \quad h(t) \le t, \tag{4.6}$$

where $h(t)$ is an arbitrary measurable function such that $t - h(t) \le \pi$ and $\alpha > 0$.

This equation has the form (3.28) where $a(t) = \alpha(|\sin t| - \sin t)$. Let us check that the conditions of Corollary 3.10 hold. It is evident that $\int_0^\infty a(s)ds = \infty$. We have

$$\limsup_{t\to\infty} \int_{h(t)}^{t} a(s)ds \le \limsup_{t\to\infty} \int_{t-\pi}^{t} a(s)ds \le -\alpha \int_{\pi}^{2\pi} 2\sin s\, ds = 4\alpha. \tag{4.7}$$

If $\alpha < 0.25(1 + 1/e)$, then condition (3.27) holds, hence all solutions of (4.6) tend to zero as $t \to \infty$.

Statements 1–3 fail for this equation. In Statements 1 and 2 the delay should be continuous. In Statement 3 function $A(t) = \int_0^t a(s)ds$ should be strictly increasing.

Consider now the general equation (1.1) with several delays. The following two statements are well known for this equation.

Statement 6 (see [6]). Suppose $a_k(t) \ge 0$, $h_k(t) \le t$ are continuous functions and

$$\limsup_{t\to\infty} a_k(t) \limsup_{t\to\infty} (t - h_k(t)) \le 1. \tag{4.8}$$

Then all solutions of (1.1) are bounded and 1 in the right-hand side of (4.8) is the best possible constant.

If $\sum_{k=1}^m a_k(t) > 0$ and the strict inequality in (4.8) is valid then all solutions of (1.1) tend to zero as $t \to \infty$.

If $a_k(t)$ are constants then in (4.8) the number 1 can be replaced by $3/2$.

Statement 7 (see [7]). Suppose $a_k(t) \ge 0$, $h_k(t) \le t$ are continuous, $h_1(t) \le h_2(t) \le \cdots \le h_m(t)$ and

$$\limsup_{t\to\infty} \int_{h_1(t)}^{t} \sum_{k=1}^{m} a_k(s)ds \le \frac{3}{2}. \tag{4.9}$$

Then any solution of (1.1) tends to a constant as $t \to \infty$.

If in addition $\int_0^\infty \sum_{k=1}^m a_k(s)ds = \infty$, then all solutions of (1.1) tend to zero as $t \to \infty$.

Example 4.2. Consider the equation

$$\dot{x}(t) + \frac{\alpha}{t}x\left(\frac{t}{2} - \sin t\right) + \frac{\beta}{t}x\left(\frac{t}{2}\right) = 0, \quad t \ge t_0 > 0, \tag{4.10}$$

where $\alpha > 0$, $\beta > 0$. Denote $p(t) = 1/t$, $h(t) = t/2 - \sin t$, $g(t) = t/2$.

to Contents, page 1

We apply Corollary 3.6. Since $\lim_{t\to\infty}[\ln(t/2) - \ln(t/2 - \sin t)] = 0$, then

$$\limsup_{t\to\infty}\left(\alpha\int_{h(t)}^{t} p(s)ds + \beta\int_{g(t)}^{t} p(s)ds\right) \le (\alpha+\beta)\ln 2. \tag{4.11}$$

Hence if $\alpha+\beta < (1/\ln 2)(1+1/e)$ then (4.10) is asymptotically stable. Statement 4 fails for this equation since the delays are unbounded. Statement 5 fails for this equation since neither $h(t) \le g(t)$ nor $g(t) \le h(t)$ holds.

Stability results where the nondelay term dominates over the delayed terms are well known beginning with the book of Krasovskii [30]. The following result is cited from the monograph [22].

Statement 8 (see [22]). Suppose $a(t)$, $b_k(t)$, $t - h_k(t)$ are bounded continuous functions, there exist δ, k, $\delta > 0$, $0 < k < 1$, such that $a(t) \ge \delta$ and $\sum_{k=1}^{m} |b_k(t)| < k\delta$. Then the equation

$$\dot{x}(t) + a(t)x(t) + \sum_{k=1}^{m} b_k(t)x(h_k(t)) = 0 \tag{4.12}$$

is exponentially stable.

In Corollary 3.13 we obtained a similar result without the assumption that the parameters of the equation are continuous functions and the delays are bounded.

Example 4.3. Consider the equation

$$\dot{x}(t) + \frac{1}{t}x(t) + \frac{\alpha}{t}x\left(\frac{t}{2}\right) = 0, \quad t \ge t_0 > 0. \tag{4.13}$$

If $\alpha < 1$ then by Corollary 3.13 all solutions of (4.13) tend to zero. The delay is unbounded, thus Statement 8 fails for this equation.

In [31] the authors considered a delay autonomous equation with linear and nonlinear parts, where the differential equation with the linear part only has a positive fundamental function and the linear part dominates over the nonlinear one. They generalized the early result of Győri [32] and some results of [33].

In Theorem 3.11 we obtained a similar result for a linear nonautonomous equation without the assumption that coefficients and delays are continuous.

In all the results of the paper we assumed that all or several coefficients of equations considered here are nonnegative. Stability results for (3.28) with oscillating coefficient $a(t)$ were obtained in [34, 35].

We conclude this paper with some open problems.

(1) Is the constant $1 + 1/e$ sharp? Prove or disprove that in Corollary 3.10 the constant $1 + 1/e$ can be replaced by the constant $3/2$.
Note that all known proofs with the constant $3/2$ apply methods which are not applicable for equations with measurable parameters.

to Contents, page 1

(2) Suppose (2.11), (2.12) hold and

$$\limsup_{t\to\infty}\sum_{k=1}^{m}\frac{|a_k(t)|}{\sum_{i=1}^{m}a_i(t)}\int_{h_k(t)}^{t}\sum_{i=1}^{m}a_i(s)ds<1+\frac{1}{e}. \tag{4.14}$$

Prove or disprove that (1.1) is exponentially stable.
The solution of this problem will improve Theorem 3.3.

(3) Suppose (1.1) is exponentially stable. Prove or disprove that the ordinary differential equation (3.41) is also exponentially (asymptotically) stable, without restrictions on the signs of coefficients $a_k(t) \geq 0$, as in Theorem 3.16. The solution of this problem would improve Theorem 3.16.

Acknowledgments

L. Berezansky is partially supported by Israeli Ministry of Absorption. E. Braverman is partially supported by the NSERC Research Grant.

References

[1] L. Berezansky and E. Braverman, "On stability of some linear and nonlinear delay differential equations," *Journal of Mathematical Analysis and Applications*, vol. 314, no. 2, pp. 391–411, 2006.
[2] L. Berezansky and E. Braverman, "On exponential stability of linear differential equations with several delays," *Journal of Mathematical Analysis and Applications*, vol. 324, no. 2, pp. 1336–1355, 2006.
[3] L. Berezansky and E. Braverman, "Explicit exponential stability conditions for linear differential equations with several delays," *Journal of Mathematical Analysis and Applications*, vol. 332, no. 1, pp. 246–264, 2007.
[4] T. Yoneyama and J. Sugie, "On the stability region of scalar delay-differential equations," *Journal of Mathematical Analysis and Applications*, vol. 134, no. 2, pp. 408–425, 1988.
[5] T. Yoneyama, "The 3/2 stability theorem for one-dimensional delay-differential equations with unbounded delay," *Journal of Mathematical Analysis and Applications*, vol. 165, no. 1, pp. 133–143, 1992.
[6] T. Krisztin, "On stability properties for one-dimensional functional-differential equations," *Funkcialaj Ekvacioj*, vol. 34, no. 2, pp. 241–256, 1991.
[7] J. W.-H. So, J. S. Yu, and M.-P. Chen, "Asymptotic stability for scalar delay differential equations," *Funkcialaj Ekvacioj*, vol. 39, no. 1, pp. 1–17, 1996.
[8] I. Győri, F. Hartung, and J. Turi, "Preservation of stability in delay equations under delay perturbations," *Journal of Mathematical Analysis and Applications*, vol. 220, no. 1, pp. 290–312, 1998.
[9] I. Győri and F. Hartung, "Stability in delay perturbed differential and difference equations," in *Topics in Functional Differential and Difference Equations (Lisbon, 1999)*, vol. 29 of *Fields Inst. Commun.*, pp. 181–194, American Mathematical Society, Providence, RI, USA, 2001.
[10] S. A. Gusarenko and A. I. Domoshnitskiĭ, "Asymptotic and oscillation properties of first-order linear scalar functional-differential equations," *Differential Equations*, vol. 25, no. 12, pp. 1480–1491, 1989.
[11] T. Wang, "Inequalities and stability for a linear scalar functional differential equation," *Journal of Mathematical Analysis and Applications*, vol. 298, no. 1, pp. 33–44, 2004.
[12] J. H. Shen and J. S. Yu, "Asymptotic behavior of solutions of neutral differential equations with positive and negative coefficients," *Journal of Mathematical Analysis and Applications*, vol. 195, no. 2, pp. 517–526, 1995.
[13] X. Wang and L. Liao, "Asymptotic behavior of solutions of neutral differential equations with positive and negative coefficients," *Journal of Mathematical Analysis and Applications*, vol. 279, no. 1, pp. 326–338, 2003.
[14] Z. Zhang and Z. Wang, "Asymptotic behavior of solutions of neutral differential equations with positive and negative coefficients," *Annals of Differential Equations*, vol. 17, no. 3, pp. 295–305, 2001.

to Contents, page 1

[15] Z. Zhang and J. Yu, "Asymptotic behavior of solutions of neutral difference equations with positive and negative coefficients," *Mathematical Sciences Research Hot-Line*, vol. 2, no. 6, pp. 1–12, 1998.
[16] N. V. Azbelev and P. M. Simonov, *Stability of Differential Equations with Aftereffect*, vol. 20 of *Stability and Control: Theory, Methods and Applications*, Taylor & Francis, London, UK, 2003.
[17] A. Ivanov, E. Liz, and S. Trofimchuk, "Halanay inequality, Yorke 3/2 stability criterion, and differential equations with maxima," *The Tohoku Mathematical Journal*, vol. 54, no. 2, pp. 277–295, 2002.
[18] E. Liz, V. Tkachenko, and S. Trofimchuk, "A global stability criterion for scalar functional differential equations," *SIAM Journal on Mathematical Analysis*, vol. 35, no. 3, pp. 596–622, 2003.
[19] X. H. Tang, "Asymptotic behavior of delay differential equations with instantaneously terms," *Journal of Mathematical Analysis and Applications*, vol. 302, no. 2, pp. 342–359, 2005.
[20] V. V. Malygina, "Some criteria for stability of equations with retarded argument," *Differential Equations*, vol. 28, no. 10, pp. 1398–1405, 1992.
[21] V. V. Malygina, "Stability of solutions of some linear differential equations with aftereffect," *Russian Mathematics*, vol. 37, no. 5, pp. 63–75, 1993.
[22] J. K. Hale and S. M. Verduyn Lunel, *Introduction to Functional-Differential Equations*, vol. 99 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 1993.
[23] N. V. Azbelev, L. Berezansky, and L. F. Rahmatullina, "A linear functional-differential equation of evolution type," *Differential Equations*, vol. 13, no. 11, pp. 1331–1339, 1977.
[24] I. Győri and G. Ladas, *Oscillation Theory of Delay Differential Equations with Applications*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, NY, USA, 1991.
[25] L. Berezansky and E. Braverman, "On non-oscillation of a scalar delay differential equation," *Dynamic Systems and Applications*, vol. 6, no. 4, pp. 567–580, 1997.
[26] L. Berezansky and E. Braverman, "Preservation of exponential stability for linear non-autonomous functional differential systems," *Automatica*, vol. 46, no. 12, pp. 2077–2081, 2010.
[27] L. Berezansky and E. Braverman, "Preservation of the exponential stability under perturbations of linear delay impulsive differential equations," *Zeitschrift für Analysis und ihre Anwendungen*, vol. 14, no. 1, pp. 157–174, 1995.
[28] G. Ladas, Y. G. Sficas, and I. P. Stavroulakis, "Asymptotic behavior of solutions of retarded differential equations," *Proceedings of the American Mathematical Society*, vol. 88, no. 2, pp. 247–253, 1983.
[29] A. D. Myshkis, *Differential Equations with Retarded Argument*, Nauka, Moscow, Russia, 1951.
[30] N. Krasovskii, *Stability of Motion*, Nauka, Moscow, Russia, 1959, translation, Stanford University Press, 1963.
[31] E. Liz and M. Pituk, "Exponential stability in a scalar functional differential equation," *Journal of Inequalities and Applications*, Article ID 37195, 10 pages, 2006.
[32] I. Győri, "Interaction between oscillations and global asymptotic stability in delay differential equations," *Differential and Integral Equations*, vol. 3, no. 1, pp. 181–200, 1990.
[33] T. Faria and W. Huang, "Special solutions for linear functional differential equations and asymptotic behaviour," *Differential and Integral Equations*, vol. 18, no. 3, pp. 337–360, 2005.
[34] I. Győri and M. Pituk, "Stability criteria for linear delay differential equations," *Differential and Integral Equations*, vol. 10, no. 5, pp. 841–852, 1997.
[35] L. Berezansky and E. Braverman, "On exponential stability of a linear delay differential equation with an oscillating coefficient," *Applied Mathematics Letters*, vol. 22, no. 12, pp. 1833–1837, 2009.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 591254, 34 pages
doi:10.1155/2011/591254

Research Article

Nonoscillation of Second-Order Dynamic Equations with Several Delays

Elena Braverman[1] and Başak Karpuz[2]

[1] *Department of Mathematics and Statistics, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada T2N 1N4*
[2] *Department of Mathematics, Faculty of Science and Arts, ANS Campus, Afyon Kocatepe University, 03200 Afyonkarahisar, Turkey*

Correspondence should be addressed to Elena Braverman, maelena@math.ucalgary.ca

Received 30 December 2010; Accepted 13 February 2011

Academic Editor: Miroslava Růžičková

Existence of nonoscillatory solutions for the second-order dynamic equation $(A_0 x^{\Delta})^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)x(\alpha_i(t)) = 0$ for $t \in [t_0,\infty)_{\mathbb{T}}$ is investigated in this paper. The results involve nonoscillation criteria in terms of relevant dynamic and generalized characteristic inequalities, comparison theorems, and explicit nonoscillation and oscillation conditions. This allows to obtain most known nonoscillation results for second-order delay differential equations in the case $A_0(t) \equiv 1$ for $t \in [t_0,\infty)_{\mathbb{R}}$ and for second-order nondelay difference equations ($\alpha_i(t) = t+1$ for $t \in [t_0,\infty)_{\mathbb{N}}$). Moreover, the general results imply new nonoscillation tests for delay differential equations with arbitrary A_0 and for second-order delay difference equations. Known nonoscillation results for quantum scales can also be deduced.

1. Introduction

This paper deals with second-order linear delay dynamic equations on time scales. Differential equations of the second order have important applications and were extensively studied; see, for example, the monographs of Agarwal et al. [1], Erbe et al. [2], Győri and Ladas [3], Ladde et al. [4], Myškis [5], Norkin [6], Swanson [7], and references therein. Difference equations of the second order describe finite difference approximations of second-order differential equations, and they also have numerous applications.

We study nonoscillation properties of these two types of equations and some of their generalizations. The main result of the paper is that under some natural assumptions for a delay dynamic equation the following four assertions are equivalent: nonoscillation of solutions of the equation on time scales and of the corresponding dynamic inequality, positivity of the fundamental function, and the existence of a nonnegative solution for a generalized Riccati inequality. The equivalence of oscillation properties of the differential equation and the corresponding differential inequality can be applied to obtain new explicit nonoscillation and oscillation conditions and also to prove some well-known results in a different way. A generalized Riccati inequality is used to compare oscillation properties of two equations without comparing their solutions. These results can be regarded as a natural generalization of the well-known Sturm-Picone comparison theorem for a second-order ordinary differential equation; see [7, Section 1.1]. Applying positivity of the fundamental function, positive solutions of two equations can be compared. There are many results of this kind for delay differential equations of the first-order and only a few for second-order equations. Myškis [5] obtained one of the first comparison theorems for second-order differential equations. The results presented here are generalizations of known nonoscillation tests even for delay differential equations (when the time scale is the real line).

The paper also contains conditions on the initial function and initial values which imply that the corresponding solution is positive. Such conditions are well known for first-order delay differential equations; however, to the best of our knowledge, the only paper concerning second-order equations is [8].

From now on, we will without furthermore mentioning suppose that the time scale $\mathbb{T}$ is unbounded from above. The purpose of the present paper is to study nonoscillation of the delay dynamic equation

$$\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)x(\alpha_i(t)) = f(t) \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}, \tag{1.1}$$

where $n \in \mathbb{N}$, $t_0 \in \mathbb{T}$, $f \in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R})$ is the forcing term, $A_0 \in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R}^+)$, and for all $i \in [1,n]_{\mathbb{N}}$, $A_i \in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R})$ is the coefficient corresponding to the function α_i, where $\alpha_i \le \sigma$ on $[t_0,\infty)_{\mathbb{T}}$.

In this paper, we follow the method employed in [8] for second-order delay differential equations. The method can also be regarded as an application of that used in [9] for first-order dynamic equations.

As a special case, the results of the present paper allow to deduce nonoscillation criteria for the delay differential equation

$$\left(A_0 x'\right)'(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)x(\alpha_i(t)) = 0 \quad \text{for } t \in [t_0,\infty)_{\mathbb{R}}, \tag{1.2}$$

in the case $A_0(t) \equiv 1$ for $t \in [t_0,\infty)_{\mathbb{R}}$, they coincide with theorems in [8]. The case of a "quickly growing" function A_0 when the integral of its reciprocal can converge is treated separately.

Let us recall some known nonoscillation and oscillation results for the ordinary differential equations

$$\left(A_0 x'\right)'(t) + A_1(t)x(t) = 0 \quad \text{for } t \in [t_0,\infty)_{\mathbb{R}}, \tag{1.3}$$

$$x''(t) + A_1(t)x(t) = 0 \quad \text{for } t \in [t_0,\infty)_{\mathbb{R}}, \tag{1.4}$$

where A_1 is nonnegative, which are particular cases of (1.2) with $n = 1$, $\alpha_1(t) = t$, and $A_0(t) \equiv 1$ for all $t \in [t_0,\infty)_{\mathbb{R}}$.

to Contents, page 1

In [10], Leighton proved the following well-known oscillation test for (1.4); see [10, 11].

Theorem A (see [10]). *Assume that*

$$\int_{t_0}^{\infty} \frac{1}{A_0(\eta)} d\eta = \infty, \qquad \int_{t_0}^{\infty} A_1(\eta) d\eta = \infty, \tag{1.5}$$

then (1.3) *is oscillatory.*

This result for (1.4) was obtained by Wintner in [12] without imposing any sign condition on the coefficient A_1.

In [13], Kneser proved the following result.

Theorem B (see [13]). *Equation* (1.4) *is nonoscillatory if* $t^2 A_1(t) \leq 1/4$ *for all* $t \in [t_0, \infty)_{\mathbb{R}}$, *while oscillatory if* $t^2 A_1(t) > \lambda_0/4$ *for all* $t \in [t_0, \infty)_{\mathbb{R}}$ *and some* $\lambda_0 \in (1, \infty)_{\mathbb{T}}$.

In [14], Hille proved the following result, which improves the one due to Kneser; see also [14–16].

Theorem C (see [14]). *Equation* (1.4) *is nonoscillatory if*

$$t \int_{t}^{\infty} A_1(\eta) d\eta \leq \frac{1}{4} \quad \forall t \in [t_0, \infty)_{\mathbb{R}}, \tag{1.6}$$

while it is oscillatory if

$$t \int_{t}^{\infty} A_1(\eta) d\eta > \frac{\lambda_0}{4} \quad \forall t \in [t_0, \infty)_{\mathbb{R}} \text{ and some } \lambda_0 \in (1, \infty)_{\mathbb{R}}. \tag{1.7}$$

Another particular case of (1.1) is the second-order delay difference equation

$$\Delta(A_0 \Delta x)(k) + \sum_{i \in [1,n]_{\mathbb{N}}} A_i(k) x(\alpha_i(k)) = 0 \quad \text{for} \in [k_0, \infty)_{\mathbb{N}}, \tag{1.8}$$

to the best of our knowledge, there are very few nonoscillation results for this equation; see, for example, [17]. However, nonoscillation properties of the nondelay equations

$$\Delta(A_0 \Delta x)(k) + A_1(k) x(k+1) = 0 \quad \text{for } k \in [k_0, \infty)_{\mathbb{N}}, \tag{1.9}$$

$$\Delta^2 x(k) + A_1(k) x(k+1) = 0 \quad \text{for } k \in [k_0, \infty)_{\mathbb{N}} \tag{1.10}$$

have been extensively studied in [1, 18–22]; see also [23]. In particular, the following result is valid.

Theorem D. *Assume that*

$$\sum_{j=k_0}^{\infty} A_1(j) = \infty, \tag{1.11}$$

then (1.10) *is oscillatory.*

The following theorem can be regarded as the discrete analogue of the nonoscillation result due to Kneser.

Theorem E. *Assume that* $k(k+1)A_1(k) \le 1/4$ *for all* $k \in [k_0, \infty)_{\mathbb{N}}$, *then* (1.10) *is nonoscillatory.*

Hille's result in [14] also has a counterpart in the discrete case. In [22], Zhou and Zhang proved the nonoscillation part, and in [24], Zhang and Cheng justified the oscillation part which generalizes Theorem E.

Theorem F (see [22, 24]). *Equation* (1.10) *is nonoscillatory if*

$$k\sum_{j=k}^{\infty} A_1(j) \le \frac{1}{4} \quad \forall k \in [k_0, \infty)_{\mathbb{N}}, \tag{1.12}$$

while is oscillatory if

$$k\sum_{j=k}^{\infty} A_1(j) > \frac{\lambda_0}{4} \quad \forall k \in [k_0, \infty)_{\mathbb{N}} \text{ and some } \lambda_0 \in (1, \infty)_{\mathbb{R}}. \tag{1.13}$$

In [23], Tang et al. studied nonoscillation and oscillation of the equation

$$\Delta^2 x(k) + A_1(k)x(k) = 0 \quad \text{for } k \in [k_0, \infty)_{\mathbb{N}}, \tag{1.14}$$

where $\{A_1(k)\}$ is a sequence of nonnegative reals and obtained the following theorem.

Theorem G (see [23]). *Equation* (1.14) *is nonoscillatory if* (1.12) *holds, while is it oscillatory if* (1.13) *holds.*

These results together with some remarks on the q-difference equations will be discussed in Section 7. The readers can find some nonoscillation results for second-order nondelay dynamic equations in the papers [20, 25–29], some of which generalize some of those mentioned above.

The paper is organized as follows. In Section 2, some auxiliary results are presented. In Section 3, the equivalence of the four above-mentioned properties is established. Section 4 is dedicated to comparison results. Section 5 includes some explicit nonoscillation and oscillation conditions. A sufficient condition for existence of a positive solution is given

to Contents, page 1

in Section 6. Section 7 involves some discussion and states open problems. Section 7 as an appendix contains a short account on the fundamentals of the time scales theory.

2. Preliminary Results

Consider the following delay dynamic equation:

$$\begin{gathered}\left(A_0 x^{\Delta}\right)^{\Delta}(t)+\sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)x(\alpha_i(t))=f(t)\quad \text{for } t\in[t_0,\infty)_{\mathbb{T}},\\ x(t_0)=x_1,\quad x^{\Delta}(t_0)=x_2,\quad x(t)=\varphi(t)\quad \text{for } t\in[t_{-1},t_0)_{\mathbb{T}},\end{gathered}\tag{2.1}$$

where $n\in\mathbb{N}$, $\mathbb{T}$ is a time scale unbounded above, $t_0\in\mathbb{T}$, $x_1,x_2\in\mathbb{R}$ are the initial values, $\varphi\in C_{\mathrm{rd}}([t_{-1},t_0)_{\mathbb{T}},\mathbb{R})$ is the initial function, such that φ has a finite left-sided limit at the initial point t_0 provided that it is left dense, $f\in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R})$ is the forcing term, $A_0\in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R}^+)$, and for all $i\in[1,n]_{\mathbb{N}}$, $A_i\in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R})$ is the coefficient corresponding to the function $\alpha_i\in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{T})$, which satisfies $\alpha_i(t)\le\sigma(t)$ for all $t\in[t_0,\infty)_{\mathbb{T}}$ and $\lim_{t\to\infty}\alpha_i(t)=\infty$. Here, we denoted

$$\alpha_{\min}(t):=\min_{i\in[1,n]_{\mathbb{N}}}\{\alpha_i(t)\}\quad \text{for } t\in[t_0,\infty)_{\mathbb{T}},\quad t_{-1}:=\inf_{t\in[t_0,\infty)_{\mathbb{T}}}\{\alpha_{\min}(t)\},\tag{2.2}$$

then t_{-1} is finite, since $\alpha_{\min}$ asymptotically tends to infinity.

Definition 2.1. A function $x:[t_{-1},\infty)_{\mathbb{T}}\to\mathbb{R}$ with $x\in C^1_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R})$ and a derivative satisfying $A_0x^{\Delta}\in C^1_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R})$ is called a *solution* of (2.1) if it satisfies the equation in the first line of (2.1) identically on $[t_0,\infty)_{\mathbb{T}}$ and also the initial conditions in the second line of (2.1).

For a given function $\varphi\in C_{\mathrm{rd}}([t_{-1},t_0)_{\mathbb{T}},\mathbb{R})$ with a finite left-sided limit at the initial point t_0 provided that it is left-dense and $x_1,x_2\in\mathbb{R}$, problem (2.1) admits a unique solution satisfying $x=\varphi$ on $[t_{-1},t_0)_{\mathbb{T}}$ with $x(t_0)=x_1$ and $x^{\Delta}(t_0)=x_2$ (see [30] and [31, Theorem 3.1]).

Definition 2.2. A solution of (2.1) is called *eventually positive* if there exists $s\in[t_0,\infty)_{\mathbb{T}}$ such that $x>0$ on $[s,\infty)_{\mathbb{T}}$, and if $(-x)$ is eventually positive, then x is called *eventually negative*. If (2.1) has a solution which is either eventually positive or eventually negative, then it is called *nonoscillatory*. A solution, which is neither eventually positive nor eventually negative, is called *oscillatory*, and (2.1) is said to be *oscillatory* provided that every solution of (2.1) is oscillatory.

For convenience in the notation and simplicity in the proofs, we suppose that functions vanish out of their specified domains, that is, let $f:D\to\mathbb{R}$ be defined for some $D\subset\mathbb{R}$, then it is always understood that $f(t)=\chi_D(t)f(t)$ for $t\in\mathbb{R}$, where χ_D is the characteristic function of the set $D\subset\mathbb{R}$ defined by $\chi_D(t)\equiv 1$ for $t\in D$ and $\chi_D(t)\equiv 0$ for $t\notin D$.

to Contents, page 1

Definition 2.3. Let $s \in \mathbb{T}$ and $s_{-1} := \inf_{t\in[s,\infty)_{\mathbb{T}}}\{\alpha_{\min}(t)\}$. The solutions $\mathcal{X}_1 = \mathcal{X}_1(\cdot, s)$ and $\mathcal{X}_2 = \mathcal{X}_2(\cdot, s)$ of the problems

$$\begin{gathered}\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)x(\alpha_i(t)) = 0 \quad \text{for } t \in [s,\infty)_{\mathbb{T}}, \\ x^{\Delta}(s) = \frac{1}{A_0(s)}, \quad x(t) \equiv 0 \quad \text{for } t \in [s_{-1}, s]_{\mathbb{T}},\end{gathered} \tag{2.3}$$

$$\begin{gathered}\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)x(\alpha_i(t)) = 0 \quad \text{for } t \in [s,\infty)_{\mathbb{T}}, \\ x^{\Delta}(s) = 0, \quad x(t) = \chi_{\{s\}}(t) \quad \text{for } t \in [s_{-1}, s]_{\mathbb{T}},\end{gathered} \tag{2.4}$$

which satisfy $\mathcal{X}_1(\cdot, s), \mathcal{X}_2(\cdot, s) \in C^1_{\mathrm{rd}}([s,\infty)_{\mathbb{T}}, \mathbb{R})$, are called the *first fundamental solution* and the *second fundamental solution* of (2.1), respectively.

The following lemma plays the major role in this paper; it presents a representation formula to solutions of (2.1) by the means of the fundamental solutions $\mathcal{X}_1$ and $\mathcal{X}_2$.

Lemma 2.4. *Let x be a solution of* (2.1), *then x can be written in the following form:*

$$x(t) = x_2\mathcal{X}_1(t,t_0) + x_1\mathcal{X}_2(t,t_0) + \int_{t_0}^{t} \mathcal{X}_1(t,\sigma(\eta))\left[f(\eta) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta)\varphi(\alpha_i(\eta))\right]\Delta\eta \tag{2.5}$$

for $t \in [t_0,\infty)_{\mathbb{T}}$.

Proof. For $t \in [t_{-1},\infty)_{\mathbb{T}}$, let

$$y(t) := \begin{cases} \displaystyle\int_{t_0}^{t} \mathcal{X}_1(t,\sigma(\eta))\left[f(\eta) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta)\varphi(\alpha_i(\eta))\right]\Delta\eta & \text{for } t \in [t_0,\infty)_{\mathbb{T}}, \\ \varphi(t) & \text{for } t \in [t_1, t_0)_{\mathbb{T}}. \end{cases} \tag{2.6}$$

We recall that $\mathcal{X}_1(\cdot, t_0)$ and $\mathcal{X}_2(\cdot, t_0)$ solve (2.3) and (2.4), respectively. To complete the proof, let us demonstrate that y solves

$$\begin{gathered}\left(A_0 y^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)y(\alpha_i(t)) = f(t) \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}, \\ y(t_0) = 0, \quad y^{\Delta}(t_0) = 0, \quad y(t) = \varphi(t) \quad \text{for } t \in [t_{-1}, t_0)_{\mathbb{T}}.\end{gathered} \tag{2.7}$$

to Contents, page 1

This will imply that the function z defined by $z := x_2\mathcal{X}_1(\cdot,t_0) + x_1\mathcal{X}_2(\cdot,t_0) + y$ on $[t_0,\infty)_{\mathbb{T}}$ is a solution of (2.1). Combining this with the uniqueness result in [31, Theorem 3.1] will complete the proof. For all $t \in [t_0,\infty)_{\mathbb{T}}$, we can compute that

$$\begin{aligned} y^{\Delta}(t) &= \int_{t_0}^{t} \mathcal{X}_1^{\Delta}(t,\sigma(\eta))\left[f(\eta) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta)\varphi(\alpha_i(\eta))\right]\Delta\eta \\ &\quad + \mathcal{X}_1(\sigma(t),\sigma(t))\left[f(t) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\varphi(\alpha_i(t))\right] \\ &= \int_{t_0}^{t} \mathcal{X}_1^{\Delta}(t,\sigma(\eta))\left[f(\eta) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta)\varphi(\alpha_i(\eta))\right]\Delta\eta. \end{aligned} \tag{2.8}$$

Therefore, $y(t_0) = 0$, $y^{\Delta}(t_0) = 0$, and $y = \varphi$ on $[t_{-1},t_0)_{\mathbb{T}}$, that is, y satisfies the initial conditions in (2.7). Differentiating y^{Δ} after multiplying by A_0 and using the properties of the first fundamental solution $\mathcal{X}_1$, we get

$$\begin{aligned} \left(A_0y^{\Delta}\right)^{\Delta}(t) &= \int_{t_0}^{t}\left(A_0\mathcal{X}_1^{\Delta}(\cdot,\sigma(\eta))\right)^{\Delta}(t)\left[f(\eta) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta)\varphi(\alpha_i(\eta))\right]\Delta\eta \\ &\quad + A_0^{\sigma}(t)\mathcal{X}_1^{\Delta}(\sigma(t),\sigma(t))\left[f(t) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\varphi(\alpha_i(t))\right] \\ &= -\sum_{j\in[1,n]_{\mathbb{N}}} A_j(t)\int_{t_0}^{\alpha_j(t)} \mathcal{X}_1(\alpha_j(t),\sigma(\eta))\left[f(\eta) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta)\varphi(\alpha_i(\eta))\right]\Delta\eta \\ &\quad - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\varphi(\alpha_i(t)) + f(t) \end{aligned} \tag{2.9}$$

for all $t \in [t_0,\infty)_{\mathbb{T}}$. For $t \in [t_0,\infty)_{\mathbb{T}}$, set $I(t) = \{i \in [1,n]_{\mathbb{N}} : \chi_{[t_0,\infty)_{\mathbb{T}}}(\alpha_i(t)) = 1\}$ and $J(t) := \{i \in [1,n]_{\mathbb{N}} : \chi_{[t_{-1},t_0)_{\mathbb{T}}}(\alpha_i(t)) = 1\}$. Making some arrangements, for all $t \in [t_0,\infty)_{\mathbb{T}}$, we find

$$\begin{aligned} \left(A_0y^{\Delta}\right)^{\Delta}(t) &= -\sum_{j\in I(t)} A_j(t)\int_{t_0}^{\alpha_j(t)} \mathcal{X}_1(\alpha_j(t),\sigma(\eta))\left[f(\eta) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta)\varphi(\alpha_i(\eta))\right]\Delta\eta \\ &\quad - \sum_{j\in J(t)} A_j(t)\int_{t_0}^{\alpha_j(t)} \mathcal{X}_1(\alpha_j(t),\sigma(\eta))\left[f(\eta) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta)\varphi(\alpha_i(\eta))\right]\Delta\eta \\ &\quad - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\varphi(\alpha_i(t)) + f(t), \end{aligned} \tag{2.10}$$

to Contents, page 1

and thus

$$\begin{aligned}\left(A_0 y^{\Delta}\right)^{\Delta}(t) &= -\sum_{j\in I(t)} A_j(t)\int_{t_0}^{\alpha_j(t)} \mathcal{X}_1(\alpha_j(t),\sigma(\eta))f(\eta)\Delta\eta - \sum_{j\in J(t)} A_i(t)\varphi(\alpha_i(t)) + f(t)\\ &= -\sum_{j\in I(t)} A_j(t)y(\alpha_j(t)) - \sum_{j\in J(t)} A_j(t)y(\alpha_j(t)) + f(t),\end{aligned} \tag{2.11}$$

which proves that y satisfies (2.7) on $[t_0,\infty)_{\mathbb{T}}$ since $I(t)\cap J(t)=\emptyset$ and $I(t)\cup J(t)=[1,n]_{\mathbb{N}}$ for each $t\in[t_0,\infty)_{\mathbb{T}}$. The proof is therefore completed. □

Next, we present a result from [9] which will be used in the proof of the main result.

Lemma 2.5 (see [9, Lemma 2.5]). *Let $t_0 \in \mathbb{T}$ and assume that K is a nonnegative Δ-integrable function defined on $\{(t,s)\in\mathbb{T}\times\mathbb{T} : t\in[t_0,\infty)_{\mathbb{T}}, s\in[t_0,t]_{\mathbb{T}}\}$. If $f,g\in \mathrm{C}_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R})$ satisfy*

$$f(t) = \int_{t_0}^{t} K(t,\eta)f(\eta)\Delta\eta + g(t) \quad \forall t\in[t_0,\infty)_{\mathbb{T}}, \tag{2.12}$$

then $g(t)\ge 0$ for all $t\in[t_0,\infty)_{\mathbb{T}}$ implies $f(t)\ge 0$ for all $t\in[t_0,\infty)_{\mathbb{T}}$.

3. Nonoscillation Criteria

Consider the delay dynamic equation

$$\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)x(\alpha_i(t)) = 0 \quad \text{for } t\in[t_0,\infty)_{\mathbb{T}} \tag{3.1}$$

and its corresponding inequalities

$$\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)x(\alpha_i(t)) \le 0 \quad \text{for } t\in[t_0,\infty)_{\mathbb{T}}, \tag{3.2}$$

$$\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)x(\alpha_i(t)) \ge 0 \quad \text{for } t\in[t_0,\infty)_{\mathbb{T}}. \tag{3.3}$$

We now prove the following result, which plays a major role throughout the paper.

Theorem 3.1. *Suppose that the following conditions hold:*

(*A*1) $A_0\in \mathrm{C}_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R}^+)$,

(*A*2) *for* $i\in[1,n]_{\mathbb{N}}$, $A_i\in \mathrm{C}_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R}_0^+)$,

(*A*3) *for* $i\in[1,n]_{\mathbb{N}}$, $\alpha_i\in \mathrm{C}_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{T})$ *satisfies* $\alpha_i(t)\le\sigma(t)$ *for all* $t\in[t_0,\infty)_{\mathbb{T}}$ *and* $\lim_{t\to\infty}\alpha_i(t)=\infty$,

to Contents, page 1

then the following conditions are equivalent:

(i) *the second-order dynamic equation* (3.1) *has a nonoscillatory solution,*

(ii) *the second-order dynamic inequality* (3.2) *has an eventually positive solution and/or* (3.3) *has an eventually negative solution,*

(iii) *there exist a sufficiently large* $t_1 \in [t_0, \infty)_{\mathbb{T}}$ *and a function* $\Lambda \in C^1_{\mathrm{rd}}([t_1, \infty)_{\mathbb{T}}, \mathbb{R})$ *with* $\Lambda/A_0 \in \mathcal{R}^+([t_1, \infty)_{\mathbb{T}}, \mathbb{R})$ *satisfying the first-order dynamic Riccati inequality*

$$\Lambda^{\Delta}(t) + \frac{1}{A_0(t)}\Lambda^{\sigma}(t)\Lambda(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\mathrm{e}_{\ominus(\Lambda/A_0)}(t, \alpha_i(t)) \le 0 \quad \forall t \in [t_1, \infty)_{\mathbb{T}}, \tag{3.4}$$

(iv) *the first fundamental solution* $\mathcal{X}_1$ *of* (3.1) *is eventually positive, that is, there exists a sufficiently large* $t_1 \in [t_0, \infty)_{\mathbb{T}}$ *such that* $\mathcal{X}_1(t, s) > 0$ *for all* $t \in (s, \infty)_{\mathbb{T}}$ *and all* $s \in [t_1, \infty)_{\mathbb{T}}$.

Proof. The proof follows the scheme: (i)⇒(ii)⇒(iii)⇒(iv)⇒(i).

(i)⇒(ii) This part is trivial, since any eventually positive solution of (3.1) satisfies (3.2) too, which indicates that its negative satisfies (3.3).

(ii)⇒(iii) Let x be an eventually positive solution of (3.2), then there exists $t_1 \in [t_0, \infty)_{\mathbb{T}}$ such that $x(t) > 0$ for all $t \in [t_1, \infty)_{\mathbb{T}}$. We may assume without loss of generality that $x(t_1) = 1$ (otherwise, we may proceed with the function $x/x(t_1)$, which is also a solution since (3.2) is linear). Let

$$\Lambda(t) := A_0(t)\frac{x^{\Delta}(t)}{x(t)} \quad \text{for } t \in [t_1, \infty)_{\mathbb{T}}, \tag{3.5}$$

then evidently $\Lambda \in C^1_{\mathrm{rd}}([t_1, \infty)_{\mathbb{T}}, \mathbb{R})$ and

$$1 + \mu(t)\frac{\Lambda(t)}{A_0(t)} = 1 + \mu(t)\frac{x^{\Delta}(t)}{x(t)} = \frac{x^{\sigma}(t)}{x(t)} > 0 \quad \forall t \in [t_1, \infty)_{\mathbb{T}}, \tag{3.6}$$

which proves that $\Lambda/A_0 \in \mathcal{R}^+([t_1, \infty)_{\mathbb{T}}, \mathbb{R})$. This implies that the exponential function $\mathrm{e}_{\Lambda/A_0}(\cdot, t_1)$ is well defined and is positive on the entire time scale $[t_1, \infty)_{\mathbb{T}}$; see [32, Theorem 2.48]. From (3.5), we see that Λ satisfies the ordinary dynamic equation

$$\begin{gathered} x^{\Delta}(t) = \frac{\Lambda(t)}{A_0(t)}x(t) \quad \text{for } t \in [t_1, \infty)_{\mathbb{T}}, \\ x(t_1) = 1, \end{gathered} \tag{3.7}$$

whose unique solution is

$$x(t) = \mathrm{e}_{\Lambda/A_0}(t, t_1) \quad \forall t \in [t_1, \infty)_{\mathbb{T}}, \tag{3.8}$$

to Contents, page 1

see [32, Theorem 2.77]. Hence, using (3.8), for all $t \in [t_1, \infty)_{\mathbb{T}}$, we get

$$\begin{aligned} x^{\Delta}(t) &= \frac{\Lambda(t)}{A_0(t)} \mathrm{e}_{\Lambda/A_0}(t, t_1), \\ \left(A_0 x^{\Delta}\right)^{\Delta}(t) &= \left(\Lambda \mathrm{e}_{\Lambda/A_0}(\cdot, t_1)\right)^{\Delta}(t) = \Lambda^{\Delta}(t) \mathrm{e}_{\Lambda/A_0}(t, t_1) + \Lambda^{\sigma}(t) \mathrm{e}^{\Delta}_{\Lambda/A_0}(t, t_1) \\ &= \Lambda^{\Delta}(t) \mathrm{e}_{\Lambda/A_0}(t, t_1) + \frac{1}{A_0(t)} \Lambda^{\sigma}(t) \Lambda(t) \mathrm{e}_{\Lambda/A_0}(t, t_1), \end{aligned} \tag{3.9}$$

which gives by substituting into (3.2) and using [32, Theorem 2.36] that

$$\begin{aligned} 0 &\geq \Lambda^{\Delta}(t) \mathrm{e}_{\Lambda/A_0}(t, t_1) + \frac{1}{A_0(t)} \Lambda^{\sigma}(t) \Lambda(t) \mathrm{e}_{\Lambda/A_0}(t, t_1) + \sum_{i \in [1,n]_{\mathbb{N}}} A_i(t) \mathrm{e}_{\Lambda/A_0}(\alpha_i(t), t_1) \\ &= \mathrm{e}_{\Lambda/A_0}(t, t_1) \left[\Lambda^{\Delta}(t) + \frac{1}{A_0(t)} \Lambda^{\sigma}(t) \Lambda(t) + \sum_{i \in [1,n]_{\mathbb{N}}} A_i(t) \frac{\mathrm{e}_{\Lambda/A_0}(\alpha_i(t), t_1)}{\mathrm{e}_{\Lambda/A_0}(t, t_1)} \right] \\ &= \mathrm{e}_{\Lambda/A_0}(t, t_1) \left[\Lambda^{\Delta}(t) + \frac{1}{A_0(t)} \Lambda^{\sigma}(t) \Lambda(t) + \sum_{i \in [1,n]_{\mathbb{N}}} A_i(t) \mathrm{e}_{\ominus(\Lambda/A_0)}(t, \alpha_i(t)) \right] \end{aligned} \tag{3.10}$$

for all $t \in [t_1, \infty)_{\mathbb{T}}$. Since the expression in the brackets is the same as the left-hand side of (3.4) and $\mathrm{e}_{\Lambda/A_0}(\cdot, t_1) > 0$ on $[t_1, \infty)_{\mathbb{T}}$, the function Λ is a solution of (3.4).

(iii)⇒(iv) Consider the initial value problem

$$\begin{gathered} \left(A_0 x^{\Delta}\right)^{\Delta}(t) + \sum_{i \in [1,n]_{\mathbb{N}}} A_i(t) x(\alpha_i(t)) = f(t) \quad \text{for } t \in [t_1, \infty)_{\mathbb{T}}, \\ x^{\Delta}(t_1) = 0, \quad x(t) \equiv 0 \quad \text{for } t \in [t_{-1}, t_1]_{\mathbb{T}}. \end{gathered} \tag{3.11}$$

Denote

$$g(t) := A_0(t) x^{\Delta}(t) - \Lambda(t) x(t) \quad \text{for } t \in [t_1, \infty)_{\mathbb{T}}, \tag{3.12}$$

where x is any solution of (3.11) and Λ is a solution of (3.4). From (3.12), we have

$$\begin{gathered} x^{\Delta}(t) = \frac{\Lambda(t)}{A_0(t)} x(t) + \frac{g(t)}{A_0(t)} \quad \text{for } t \in [t_1, \infty)_{\mathbb{T}}, \\ x(t_1) = 0, \end{gathered} \tag{3.13}$$

whose unique solution is

$$x(t) = \int_{t_1}^{t} \mathrm{e}_{\Lambda/A_0}(t, \sigma(\eta)) \frac{g(\eta)}{A_0(\eta)} \Delta\eta \quad \forall t \in [t_1, \infty)_{\mathbb{T}}, \tag{3.14}$$

to Contents, page 1

see [32, Theorem 2.77]. Now, for all $t \in [t_1, \infty)_{\mathbb{T}}$, we compute that

$$\begin{aligned} x(t) &= \mathrm{e}_{\ominus(\Lambda/A_0)}(\sigma(t), t)\left[\int_{t_1}^{\sigma(t)} \mathrm{e}_{\Lambda/A_0}(\sigma(t), \sigma(\eta))\frac{g(\eta)}{A(\eta)}\Delta\eta - \mu(t)\mathrm{e}_{\Lambda/A_0}(\sigma(t), \sigma(t))\frac{g(t)}{A_0(t)}\right] \\ &= \frac{A_0(t)}{A_0(t)+\mu(t)\Lambda(t)}\left[x^{\sigma}(t) - \mu(t)\frac{g(t)}{A_0(t)}\right] \\ &= \frac{1}{A_0(t)+\mu(t)\Lambda(t)}\left[A_0(t)x^{\sigma}(t) - \mu(t)g(t)\right], \end{aligned} \tag{3.15}$$

and similarly

$$\begin{aligned} x(\alpha_i(t)) &= \mathrm{e}_{\ominus(\Lambda/A_0)}(\sigma(t), \alpha_i(t)) \\ &\quad \times \left[\int_{t_1}^{\sigma(t)} \mathrm{e}_{\Lambda/A_0}(\sigma(t), \sigma(\eta))\frac{g(\eta)}{A(\eta)}\Delta\eta - \int_{\alpha_i(t)}^{\sigma(t)} \mathrm{e}_{\Lambda/A_0}(\sigma(t), \sigma(\eta))\frac{g(\eta)}{A_0(\eta)}\Delta\eta\right] \\ &= \mathrm{e}_{\ominus(\Lambda/A_0)}(\sigma(t), \alpha_i(t))\left[x^{\sigma}(t) - \int_{\alpha_i(t)}^{\sigma(t)} \mathrm{e}_{\Lambda/A_0}(\sigma(t), \sigma(\eta))\frac{g(\eta)}{A_0(\eta)}\Delta\eta\right] \\ &= \mathrm{e}_{\ominus(\Lambda/A_0)}(\sigma(t), \alpha_i(t))x^{\sigma}(t) - \int_{\alpha_i(t)}^{\sigma(t)} \mathrm{e}_{\Lambda/A_0}(\alpha_i(t), \sigma(\eta))\frac{g(\eta)}{A_0(\eta)}\Delta\eta \end{aligned} \tag{3.16}$$

for $i \in [1, n]_{\mathbb{N}}$. From (3.12) and (3.15), we have

$$\begin{aligned} \left(A_0 x^{\Delta}\right)^{\Delta}(t) &= (\Lambda x + g)^{\Delta}(t) = \Lambda^{\Delta}(t)x^{\sigma}(t) + \Lambda(t)x^{\Delta}(t) + g^{\Delta}(t) \\ &= \Lambda^{\Delta}(t)x^{\sigma}(t) + \frac{\Lambda^2(t)}{A_0(t)}x(t) + \frac{\Lambda(t)}{A_0(t)}g(t) + g^{\Delta}(t) \end{aligned} \tag{3.17}$$

for all $t \in [t_1, \infty)_{\mathbb{T}}$. We substitute (3.14), (3.15), (3.16), and (3.17) into (3.11) and find that

$$\begin{aligned} f(t) &= \left[\Lambda^{\Delta}(t)x^{\sigma}(t) + \frac{\Lambda^2(t)}{A_0(t)}x(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)x(\alpha_i(t))\right] + \frac{\Lambda(t)}{A_0(t)}g(t) + g^{\Delta}(t) \\ &= \left[\Lambda^{\Delta}(t) + \frac{\Lambda^2(t)}{A_0(t)+\mu(t)\Lambda(t)} + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\mathrm{e}_{\ominus(\Lambda/A_0)}(\sigma(t), \alpha_i(t))\right]x^{\sigma}(t) \end{aligned}$$

to Contents, page 1

$$
\begin{aligned}
&- \left[\frac{\mu(t)\Lambda^2(t)}{A_0(t)(A_0(t)+\mu(t)\Lambda(t))} g(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) \int_{\alpha_i(t)}^{\sigma(t)} \mathrm{e}_{\Lambda/A_0}(\alpha_i(t),\sigma(\eta)) \frac{g(\eta)}{A_0(\eta)} \Delta\eta \right] \\
&+ \frac{\Lambda(t)}{A_0(t)} g(t) + g^{\Delta}(t) \\
= &\left[\Lambda^{\Delta}(t) + \frac{\Lambda^2(t)}{A_0(t)+\mu(t)\Lambda(t)} + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) \mathrm{e}_{\ominus(\Lambda/A_0)}(\sigma(t),\alpha_i(t)) \right] \\
&\times \left[1 + \mu(t)\frac{\Lambda(t)}{A_0(t)} \right] \int_{t_1}^{\sigma(t)} \mathrm{e}_{\Lambda/A_0}(t,\sigma(\eta)) \frac{g(\eta)}{A_0(\eta)} \Delta\eta \\
&- \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) \int_{\alpha_i(t)}^{\sigma(t)} \mathrm{e}_{\Lambda/A_0}(\alpha_i(t),\sigma(\eta)) \frac{g(\eta)}{A_0(\eta)} \Delta\eta \\
&+ \frac{\Lambda(t)}{A_0(t)+\mu(t)\Lambda(t)} g(t) + g^{\Delta}(t)
\end{aligned}
\tag{3.18}
$$

for all $t \in [t_1,\infty)_{\mathbb{T}}$. Then, (3.18) can be rewritten as

$$
\begin{aligned}
g^{\Delta}(t) = &-\frac{\Lambda(t)}{A_0(t)+\mu(t)\Lambda(t)} g(t) + \Upsilon(t) \int_{t_1}^{\sigma(t)} \mathrm{e}_{\Lambda/A_0}(t,\sigma(\eta)) \frac{g(\eta)}{A_0(\eta)} \Delta\eta \\
&+ \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) \int_{\alpha_i(t)}^{\sigma(t)} \mathrm{e}_{\Lambda/A_0}(\alpha_i(t),\sigma(\eta)) \frac{g(\eta)}{A_0(\eta)} \Delta\eta + f(t)
\end{aligned}
\tag{3.19}
$$

for all $t \in [t_1,\infty)_{\mathbb{T}}$, where

$$
\Upsilon(t) := -\left[1 + \mu(t)\frac{\Lambda(t)}{A_0(t)} \right] \left[\Lambda^{\Delta}(t) + \frac{\Lambda^2(t)}{A_0(t)+\mu(t)\Lambda(t)} + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) \mathrm{e}_{\ominus(\Lambda/A_0)}(\sigma(t),\alpha_i(t)) \right]
\tag{3.20}
$$

for $t \in [t_1,\infty)_{\mathbb{T}}$. We now show that $\Upsilon \geq 0$ on $[t_1,\infty)_{\mathbb{T}}$. Indeed, by using (3.4) and the simple useful formula (A.2), we get

$$
\begin{aligned}
\Upsilon(t) &= -\left[\left(1 + \mu(t)\frac{\Lambda(t)}{A_0(t)} \right) \Lambda^{\Delta}(t) + \frac{1}{A_0(t)} \Lambda^2(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) \mathrm{e}_{\ominus(\Lambda/A_0)}(t,\alpha_i(t)) \right] \\
&= -\left[\Lambda^{\Delta}(t) + \frac{1}{A_0(t)} \Lambda^{\sigma}(t)\Lambda(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) \mathrm{e}_{\ominus(\Lambda/A_0)}(t,\alpha_i(t)) \right] \geq 0
\end{aligned}
\tag{3.21}
$$

to Contents, page 1

for all $t \in [t_1, \infty)_{\mathbb{T}}$. On the other hand, from (3.11) and (3.12), we see that $g(t_1) = 0$. Then, by [32, Theorem 2.77], we can write (3.19) in the equivalent form

$$g = \mathscr{A}g + h \quad on \ [t_1, \infty)_{\mathbb{T}}, \tag{3.22}$$

where, for $t \in [t_1, \infty)_{\mathbb{T}}$, we have defined

$$\begin{aligned}(\mathscr{A}g)(t) := \int_{t_1}^{t} \mathrm{e}_{-\Lambda/(A_0+\mu\Lambda)}(t, \sigma(\eta)) \Bigg[\Upsilon(\eta) \int_{t_1}^{\sigma(\eta)} \mathrm{e}_{\Lambda/A_0}(\sigma(\eta), \sigma(\zeta)) \frac{g(\zeta)}{A_0(\zeta)} \Delta\zeta \\ + \sum_{i \in [1,n]_{\mathbb{N}}} A_i(\eta) \int_{\alpha_i(\eta)}^{\sigma(\eta)} \mathrm{e}_{\Lambda/A_0}(\alpha_i(\eta), \sigma(\zeta)) \frac{g(\zeta)}{A_0(\zeta)} \Delta\zeta \Bigg] \Delta\eta,\end{aligned} \tag{3.23}$$

$$h(t) := \int_{t_1}^{t} \mathrm{e}_{\Lambda/A_0}(t, \sigma(\zeta)) f(\eta) \Delta\eta. \tag{3.24}$$

Note that $\Lambda/A_0 \in \mathcal{R}^+([t_1, \infty)_{\mathbb{T}}, \mathbb{R})$ implies $-\Lambda/(A_0 + \mu\Lambda) \in \mathcal{R}^+([t_1, \infty)_{\mathbb{T}}, \mathbb{R})$ (indeed, we have $1 - \mu\Lambda/(A_0 + \mu\Lambda) = A_0/(A_0 + \mu\Lambda) > 0$ on $[t_1, \infty)_{\mathbb{T}}$), and thus the exponential function $\mathrm{e}_{\ominus(\Lambda/A_0)}(\cdot, t_1)$ is also well defined and positive on the entire time scale $[t_1, \infty)_{\mathbb{T}}$, see [32, Exercise 2.28]. Thus, $f \geq 0$ on $[t_1, \infty)_{\mathbb{T}}$ implies $h \geq 0$ on $[t_1, \infty)_{\mathbb{T}}$. For simplicity of notation, for $s, t \in [t_1, \infty)_{\mathbb{T}}$, we let

$$\begin{aligned}K_1(t, s) := \frac{1}{A_0(s)} \int_{s}^{t} \mathrm{e}_{-\Lambda/(A_0+\mu\Lambda)}(t, \sigma(\eta)) \Upsilon(\eta) \mathrm{e}_{\Lambda/A_0}(\sigma(\eta), \sigma(s)) \Delta\eta, \\ K_2(t, s) := \frac{1}{A_0(s)} \int_{s}^{t} \mathrm{e}_{-\Lambda/(A_0+\mu\Lambda)}(t, \sigma(\eta)) \sum_{i \in [1,n]_{\mathbb{N}}} A_i(\eta) \chi_{[\alpha_i(\eta), \infty)_{\mathbb{T}}}(s) \mathrm{e}_{\Lambda/A_0}(\sigma(\eta), \sigma(s)) \Delta\eta.\end{aligned} \tag{3.25}$$

Using the change of integration order formula in [33, Lemma 1], for all $t \in [t_1, \infty)_{\mathbb{T}}$, we obtain

$$\begin{aligned}&\int_{t_1}^{t} \int_{t_1}^{\sigma(\eta)} \mathrm{e}_{-\Lambda/(A_0+\mu\Lambda)}(t, \sigma(\eta)) \Upsilon(\eta) \mathrm{e}_{\Lambda/A_0}(\sigma(\eta), \sigma(\zeta)) \frac{g(\zeta)}{A_0(\zeta)} \Delta\zeta\Delta\eta \\ &\quad = \int_{t_1}^{t} \int_{\zeta}^{t} \mathrm{e}_{-\Lambda/(A_0+\mu\Lambda)}(t, \sigma(\eta)) \Upsilon(\eta) \mathrm{e}_{\Lambda/A_0}(\sigma(\eta), \sigma(\zeta)) \frac{g(\zeta)}{A_0(\zeta)} \Delta\eta\Delta\zeta \\ &\quad = \int_{t_1}^{t} K_1(t, \zeta) g(\zeta) \Delta\zeta,\end{aligned} \tag{3.26}$$

to Contents, page 1

and similarly

$$\begin{aligned}&\int_{t_1}^{t}\int_{t_1}^{\sigma(\eta)} \mathrm{e}_{-\Lambda/(A_0+\mu\Lambda)}(t,\sigma(\eta))\sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta)\chi_{[\alpha_i(\eta),\infty)_{\mathbb{T}}}(\zeta)\mathrm{e}_{\Lambda/A_0}(\sigma(\eta),\sigma(\zeta))\frac{g(\zeta)}{A_0(\zeta)}\Delta\zeta\Delta\eta \\ &\quad=\int_{t_1}^{t} K_2(t,\zeta)g(\zeta)\Delta\zeta.\end{aligned}\tag{3.27}$$

Therefore, we can rewrite (3.23) in the equivalent form of the integral operator

$$(\mathscr{H}g)(t)=\int_{t_1}^{t}[K_1(t,\eta)+K_2(t,\eta)]g(\eta)\Delta\eta \quad \text{for } t\in[t_1,\infty)_{\mathbb{T}},\tag{3.28}$$

whose kernel is nonnegative. Consequently, using (3.22), (3.24), and (3.28), we obtain that $f\geq 0$ on $[t_1,\infty)_{\mathbb{T}}$ implies $h\geq 0$ on $[t_1,\infty)_{\mathbb{T}}$; this and Lemma 2.5 yield that $g\geq 0$ on $[t_1,\infty)_{\mathbb{T}}$. Therefore, from (3.14), we infer that if $f\geq 0$ on $[t_1,\infty)_{\mathbb{T}}$, then $x\geq 0$ on $[t_1,\infty)_{\mathbb{T}}$ too. On the other hand, by Lemma 2.4, x has the following representation:

$$x(t)=\int_{t_1}^{t}\mathcal{X}_1(t,\sigma(\eta))f(\eta)\Delta\eta \quad \text{for } t\in[t_1,\infty)_{\mathbb{T}}.\tag{3.29}$$

Since x is eventually nonnegative for any eventually nonnegative function f, we infer that the kernel $\mathcal{X}_1$ of the integral on the right-hand side of (3.29) is eventually nonnegative. Indeed, assume to the contrary that $x\geq 0$ on $[t_1,\infty)_{\mathbb{T}}$ but $\mathcal{X}_1$ is not nonnegative, then we may pick $t_2\in[t_1,\infty)_{\mathbb{T}}$ and find $s\in[t_1,t_2)_{\mathbb{T}}$ such that $\mathcal{X}_1(t_2,\sigma(s))<0$. Then, letting $f(t):=-\min\{\mathcal{X}_1(t_2,\sigma(t)),0\}\geq 0$ for $t\in[t_1,\infty)_{\mathbb{T}}$, we are led to the contradiction $x(t_2)<0$, where x is defined by (3.29). To prove that $\mathcal{X}_1$ is eventually positive, set $x(t):=\mathcal{X}_1(t,s)$ for $t\in[t_0,\infty)_{\mathbb{T}}$, where $s\in[t_1,\infty)_{\mathbb{T}}$, to see that $x\geq 0$ and $(A_0x^{\Delta})^{\Delta}\leq 0$ on $[s,\infty)_{\mathbb{T}}$, which implies A_0x^{Δ} is nonincreasing on $[s,\infty)_{\mathbb{T}}$. So that, we may let $t_1\in[t_0,\infty)_{\mathbb{T}}$ so large that x^{Δ} (i.e., A_0x^{Δ}) is of fixed sign on $[s,\infty)_{\mathbb{T}}\subset[t_1,\infty)_{\mathbb{T}}$. The initial condition and (A1) together with $x^{\Delta}(s)=1/A_0(s)>0$ imply that $x^{\Delta}>0$ on $[s,\infty)_{\mathbb{T}}$. Consequently, we have $x(t)=\mathcal{X}_1(t,s)>\mathcal{X}_1(s,s)=0$ for all $t\in(s,\infty)_{\mathbb{T}}\subset[t_1,\infty)_{\mathbb{T}}$.

(iv)⇒(i) Clearly, $\mathcal{X}_1(\cdot,t_0)$ is an eventually positive solution of (3.1).

The proof is completed. □

Let us introduce the following condition:

$(A4)$ $A_0\in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R}^+)$ with

$$\int_{t_0}^{\infty}\frac{1}{A_0(\eta)}\Delta\eta=\infty.\tag{3.30}$$

to Contents, page 1

Remark 3.2. It is well known that (A4) ensures existence of $t_1 \in [t_0,\infty)_{\mathbb{T}}$ such that $x(t)x^{\Delta}(t) \geq 0$ for all $t \in [t_1,\infty)_{\mathbb{T}}$, for any nonoscillatory solution x of (3.1). This fact follows from the formula

$$x(t) = x(s) + A_0(s)x^{\Delta}(s)\int_s^t \frac{1}{A_0(\eta)}\Delta\eta - \int_s^t \frac{1}{A_0(\eta)}\left[\int_s^{\eta} \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\zeta)x(\alpha_i(\zeta))\Delta\zeta\right]\Delta\eta \qquad (3.31)$$

for all $t \in [t_0,\infty)_{\mathbb{T}}$, obtained by integrating (3.1) twice, where $s \in [t_0,\infty)_{\mathbb{T}}$. In the case when (A4) holds, (iii) of Theorem 3.1 can be assumed to hold with $\Lambda \in C^1_{\mathrm{rd}}([t_1,\infty)_{\mathbb{T}},\mathbb{R}^+_0)$, which means that any positive (negative) solution is nondecreasing (nonincreasing).

Remark 3.3. Let (A4) hold and exist $t_1 \in [t_0,\infty)_{\mathbb{T}}$ and the function $\Lambda \in C^1_{\mathrm{rd}}([t_1,\infty)_{\mathbb{T}},\mathbb{R}^+_0)$ satisfying inequality (3.4), then the assertions (i), (iii), and (iv) of Theorem 3.1 are also valid on $[t_1,\infty)_{\mathbb{T}}$.

Remark 3.4. It should be noted that (3.4) is also equivalent to the inequality

$$\Lambda^{\Delta}(t) + \frac{\Lambda^2(t)}{A_0(t)+\mu(t)\Lambda(t)} + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\mathrm{e}_{\ominus(\Lambda/A_0)}(\sigma(t),\alpha_i(t)) \leq 0 \quad \forall t \in [t_1,\infty)_{\mathbb{T}}, \qquad (3.32)$$

see (3.20) and compare with [26, 28, 29, 34].

Example 3.5. For $\mathbb{T} = \mathbb{R}$, (3.4) has the form

$$\Lambda'(t) + \frac{1}{A_0(t)}\Lambda^2(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\exp\left\{-\int_{\alpha_i(t)}^t \frac{\Lambda(\eta)}{A_0(\eta)}\mathrm{d}\eta\right\} \leq 0 \quad \forall t \in [t_1,\infty)_{\mathbb{R}}, \qquad (3.33)$$

see [8] for the case $A_0(t) \equiv 1,\ t \in [t_0,\infty)_{\mathbb{R}}$, and [35] for $n = 1,\ \alpha_1(t) = t,\ t \in [t_0,\infty)_{\mathbb{R}}$.

Example 3.6. For $\mathbb{T} = \mathbb{N}$, (3.4) becomes

$$\Delta\Lambda(k) + \frac{\Lambda^2(k)}{A_0(k)+\Lambda(k)} + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(k)\prod_{j=\alpha_i(k)}^{k} \frac{A_0(j)}{A_0(j)+\Lambda(j)} \leq 0 \quad \forall k \in [k_1,\infty)_{\mathbb{N}}, \qquad (3.34)$$

where the product over the empty set is assumed to be equal to one; see [1, 18] (or (1.10)) for $n = 1$, $\alpha_1(k) = k+1$, $k \in [k_0,\infty)_{\mathbb{N}}$, and [20] for $n = 1$, $A_0(k) \equiv 1$, $\alpha_1(k) = k+1$, $k \in [k_0,\infty)_{\mathbb{N}}$. It should be mentioned that in the literature all the results relating difference equations with discrete Riccati equations consider only the nondelay case. This result in the discrete case is therefore new.

to Contents, page 1

Example 3.7. For $\mathbb{T} = \overline{q^{\mathbb{Z}}}$ with $q \in (1,\infty)_{\mathbb{R}}$, under the same assumption on the product as in the previous example, condition (3.4) reduces to the inequality

$$D_q\Lambda(t) + \frac{\Lambda^2(t)}{A_0(t) + (q-1)t\Lambda(t)} + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) \prod_{\eta=\log_q(\alpha_i(t))}^{\log_q(t)} \frac{A_0(q^\eta)}{A_0(q^\eta) + (q-1)q^\eta\Lambda(q^\eta)} \le 0 \quad (3.35)$$

for all $t \in [t_1,\infty)_{\overline{q^{\mathbb{Z}}}}$.

4. Comparison Theorems

Theorem 3.1 can be employed to obtain comparison nonoscillation results. To this end, together with (3.1), we consider the second-order dynamic equation

$$\left(A_0 x^\Delta\right)^\Delta(t) + \sum_{i\in[1,n]_{\mathbb{N}}} B_i(t)x(\alpha_i(t)) = 0 \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}, \quad (4.1)$$

where $B_i \in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R})$ for $i \in [1,n]_{\mathbb{N}}$.

The following theorem establishes the relation between the first fundamental solution of the model equation with positive coefficients and comparison (4.1) with coefficients of arbitrary signs.

Theorem 4.1. *Suppose that (A2), (A3), (A4), and the following condition hold:*

(A5) *for* $i \in [1,n]_{\mathbb{N}}$, $B_i \in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R})$ *with* $A_i(t) \ge B_i(t)$ *for all* $t \in [t_0,\infty)_{\mathbb{T}}$.

Assume further that (3.4) *admits a solution* $\Lambda \in C^1_{\mathrm{rd}}([t_1,\infty)_{\mathbb{T}},\mathbb{R}^+_0)$ *for some* $t_1 \in [t_0,\infty)_{\mathbb{T}}$, *then the first fundamental solution* $\mathcal{Y}_1$ *of* (4.1) *satisfies* $\mathcal{Y}_1(t,s) \ge \mathcal{X}_1(t,s) > 0$ *for all* $t \in (s,\infty)_{\mathbb{T}}$ *and all* $s \in [t_1,\infty)_{\mathbb{T}}$, *where* $\mathcal{X}_1$ *denotes the first fundamental solution of* (3.1).

Proof. We consider the initial value problem

$$\begin{gathered}\left(A_0 x^\Delta\right)^\Delta(t) + \sum_{i\in[1,n]_{\mathbb{N}}} B_i(t)x(\alpha_i(t)) = f(t) \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}, \\ x^\Delta(t_0) = 0, \quad x(t) \equiv 0 \quad \text{for } t \in [t_{-1},t_0]_{\mathbb{T}},\end{gathered} \quad (4.2)$$

where $f \in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R})$. Let $g \in C_{\mathrm{rd}}([t_1,\infty)_{\mathbb{T}},\mathbb{R})$, and define the function x as

$$x(t) = \int_{t_1}^{t} \mathcal{X}_1(t,\sigma(\eta))g(\eta)\Delta\eta \quad \forall t \in [t_1,\infty)_{\mathbb{T}}. \quad (4.3)$$

to Contents, page 1

By the Leibnitz rule (see [32, Theorem 1.117]), for all $t \in [t_1, \infty)_{\mathbb{T}}$, we have

$$x^{\Delta}(t) = \int_{t_1}^{t} \mathcal{X}_1^{\Delta}(t, \sigma(\eta)) g(\eta) \Delta\eta, \tag{4.4}$$

$$\left(A_0 x^{\Delta}\right)^{\Delta}(t) = \int_{t_1}^{t} \left(A_0 \mathcal{X}_1^{\Delta}(\cdot, \sigma(\eta))\right)^{\Delta}(t) g(\eta) \Delta\eta + g(t). \tag{4.5}$$

Substituting (4.3) and (4.5) into (4.2), we get

$$\begin{aligned} f(t) &= \int_{t_1}^{t} \left(A_0 \mathcal{X}_1^{\Delta}(\cdot, \sigma(\eta))\right)^{\Delta}(t) g(\eta) \Delta\eta + \sum_{i \in [1,n]_{\mathbb{N}}} B_i(t) \int_{t_1}^{\alpha_i(t)} \mathcal{X}_1(\alpha_i(t), \sigma(\eta)) g(\eta) \Delta\eta + g(t) \\ &= \sum_{i \in [1,n]_{\mathbb{N}}} [B_i(t) - A_i(t)] \int_{t_1}^{\alpha_i(t)} \mathcal{X}_1(\alpha_i(t), \sigma(\eta)) g(\eta) \Delta\eta + g(t) \\ &= \sum_{i \in [1,n]_{\mathbb{N}}} [B_i(t) - A_i(t)] \int_{t_1}^{t} \mathcal{X}_1(\alpha_i(t), \sigma(\eta)) g(\eta) \Delta\eta + g(t), \end{aligned} \tag{4.6}$$

where in the last step, we have used the fact that $\mathcal{X}_1(t, \sigma(s)) \equiv 0$ for all $t \in [t_1, \infty)_{\mathbb{T}}$ and all $s \in [t, \infty)_{\mathbb{T}}$. Therefore, we obtain the operator equation

$$g = \mathcal{A} g + f \quad \text{on } [t_1, \infty)_{\mathbb{T}}, \tag{4.7}$$

where

$$(\mathcal{A} g)(t) := \int_{t_1}^{t} \sum_{i \in [1,n]_{\mathbb{N}}} \mathcal{X}_1(\alpha_i(t), \sigma(\eta)) [A_i(t) - B_i(t)] g(\eta) \Delta\eta \quad \text{for } t \in [t_1, \infty)_{\mathbb{T}}, \tag{4.8}$$

whose kernel is nonnegative. An application of Lemma 2.5 shows that nonnegativity of f implies the same for g, and thus x is nonnegative by (4.3). On the other hand, by Lemma 2.4, x has the representation

$$x(t) = \int_{t_0}^{t} \mathcal{Y}_1(t, \sigma(\eta)) f(\eta) \Delta\eta \quad \forall t \in [t_0, \infty)_{\mathbb{T}}. \tag{4.9}$$

Proceeding as in the proof of the part (iii)⇒(iv) of Theorem 3.1, we conclude that the first fundamental solution $\mathcal{Y}_1$ of (4.1) satisfies $\mathcal{Y}_1(t, s) \geq 0$ for all $t \in (s, \infty)_{\mathbb{T}}$ and all $s \in [t_1, \infty)_{\mathbb{T}}$. To complete the proof, we have to show that $\mathcal{Y}_1(t, s) \geq \mathcal{X}_1(t, s) > 0$ for all $t \in (s, \infty)_{\mathbb{T}}$ and all $s \in [t_1, \infty)_{\mathbb{T}}$. Clearly, for any fixed $s \in [t_1, \infty)_{\mathbb{T}}$ and all $t \in [s, \infty)_{\mathbb{T}}$, we have

$$\left(A_0 \mathcal{Y}_1^{\Delta}(\cdot, s)\right)^{\Delta}(t) + \sum_{i \in [1,n]_{\mathbb{N}}} A_i(t) \mathcal{Y}_1(\alpha_i(t), s) = \sum_{i \in [1,n]_{\mathbb{N}}} [A_i(t) - B_i(t)] \mathcal{Y}_1(\alpha_i(t), s), \tag{4.10}$$

to Contents, page 1

which by the solution representation formula yields that

$$\mathcal{Y}_1(t,s) = \mathcal{X}_1(t,s) + \int_s^t \mathcal{X}_1(t,\sigma(\eta)) \sum_{i\in[1,n]_{\mathbb{N}}} [A_i(\eta) - B_i(\eta)]\mathcal{Y}_1(\alpha_i(\eta),s)\Delta\eta \geq \mathcal{X}_1(t,s) \quad (4.11)$$

for all $t \in [s,\infty)_{\mathbb{T}}$. This completes the proof since the first fundamental solution $\mathcal{X}_1$ satisfies $\mathcal{X}_1(t,s) > 0$ for all $t \in (s,\infty)_{\mathbb{T}}$ and all $s \in [t_1,\infty)_{\mathbb{T}}$ by Remark 3.3. □

Corollary 4.2. *Suppose that (A1), (A2), (A3), and (A5) hold, and* (3.1) *has a nonoscillatory solution on* $[t_1,\infty)_{\mathbb{T}} \subset [t_0,\infty)_{\mathbb{T}}$, *then* (4.1) *admits a nonoscillatory solution on* $[t_2,\infty)_{\mathbb{T}} \subset [t_1,\infty)_{\mathbb{T}}$.

Corollary 4.3. *Assume that (A2) and (A3) hold.*

(i) *If (A1) holds and the dynamic inequality*

$$\left(A_0x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i^{+}(t)x(\alpha_i(t)) \leq 0 \quad \textit{for } t \in [t_0,\infty)_{\mathbb{T}}, \quad (4.12)$$

where $A_i^{+}(t) := \max\{A_i(t),0\}$ *for* $t \in [t_0,\infty)_{\mathbb{T}}$ *and* $i \in [1,n]_{\mathbb{N}}$, *has a positive solution on* $[t_0,\infty)_{\mathbb{T}}$, *then* (3.1) *also admits a positive solution on* $[t_1,\infty)_{\mathbb{T}} \subset [t_0,\infty)_{\mathbb{T}}$.

(ii) *If (A4) holds and there exist a sufficiently large* $t_1 \in [t_0,\infty)_{\mathbb{T}}$ *and a function* $\Lambda \in C^1_{\mathrm{rd}}([t_1,\infty)_{\mathbb{T}},\mathbb{R}^+_0)$ *satisfying the inequality*

$$\Lambda^{\Delta}(t) + \frac{1}{A_0(t)}\Lambda^{\sigma}(t)\Lambda(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i^{+}(t)\mathrm{e}_{\ominus(\Lambda/A_0)}(t,\alpha_i(t)) \leq 0 \quad \forall t \in [t_1,\infty)_{\mathbb{T}}, \quad (4.13)$$

then the first fundamental solution $\mathcal{X}_1$ *of* (3.1) *satisfies* $\mathcal{X}_1(t,s) > 0$ *for all* $t \in (s,\infty)_{\mathbb{T}}$ *and all* $s \in [t_1,\infty)_{\mathbb{T}}$.

Proof. Consider the dynamic equation

$$\left(A_0x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i^{+}(t)x(\alpha_i(t)) = 0 \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}. \quad (4.14)$$

Theorem 3.1 implies that for this equation the assertions (i) and (ii) hold. Since for all $i \in [1,n]_{\mathbb{N}}$, we have $A_i(t) \leq A_i^{+}(t)$ for all $t \in [t_0,\infty)_{\mathbb{T}}$, the application of Corollary 4.2 and Theorem 4.1 completes the proof. □

Now, let us compare the solutions of problem (2.1) and the following initial value problem:

$$\begin{aligned} &\left(A_0x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} B_i(t)x(\alpha_i(t)) = g(t) \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}, \\ &x(t_0) = y_1, \quad x^{\Delta}(t_0) = y_2, \quad x(t) = \varphi(t) \quad \text{for } t \in [t_{-1},t_0)_{\mathbb{T}}, \end{aligned} \quad (4.15)$$

to Contents, page 1

where $y_1, y_2 \in \mathbb{R}$ are the initial values, $\psi \in C_{\mathrm{rd}}([t_{-1}, t_0)_{\mathbb{T}}, \mathbb{R})$ is the initial function such that ψ has a finite left-sided limit at the initial point t_0 provided that it is left dense, $g \in C_{\mathrm{rd}}([t_0, \infty)_{\mathbb{T}}, \mathbb{R})$ is the forcing term.

Theorem 4.4. *Suppose that (A2), (A3), (A4), (A5), and the following condition hold:*

($A6$) *$f, g \in C_{\mathrm{rd}}([t_0, \infty)_{\mathbb{T}}, \mathbb{R})$ and $\varphi, \psi \in C_{\mathrm{rd}}([t_{-1}, t_0)_{\mathbb{T}}, \mathbb{R})$ satisfy*

$$f(t) - \sum_{i \in [1,n]_{\mathbb{N}}} B_i(t)\varphi(\alpha_i(t)) \leq g(t) - \sum_{i \in [1,n]_{\mathbb{N}}} B_i(t)\psi(\alpha_i(t)) \quad \forall t \in [t_0, \infty)_{\mathbb{T}}. \tag{4.16}$$

Moreover, let (2.1) *have a positive solution x on $[t_0, \infty)_{\mathbb{T}}$, $y_1 = x_1$, and $y_2 \geq x_2$, then the solution y of* (4.15) *satisfies $y(t) \geq x(t)$ for all $t \in [t_0, \infty)_{\mathbb{T}}$.*

Proof. By Theorem 3.1 and Remark 3.3, we can assume that $\Lambda \in C_{\mathrm{rd}}([t_0, \infty)_{\mathbb{T}}, \mathbb{R}_0^+)$ is a solution of the dynamic Riccati inequality (3.4), then by ($A5$), the function Λ is also a solution of the dynamic Riccati inequality

$$\Lambda^{\Delta}(t) + \frac{1}{A_0(t)}\Lambda^{\sigma}(t)\Lambda(t) + \sum_{i \in [1,n]_{\mathbb{N}}} B_i(t) \mathrm{e}_{\ominus(\Lambda/A_0)}(t, \alpha_i(t)) \leq 0 \quad \forall t \in [t_0, \infty)_{\mathbb{T}}, \tag{4.17}$$

which is associated with (4.15). Hence, by Theorem 3.1 and Remark 3.3, the first fundamental solution $\mathcal{Y}_1$ of (4.15) satisfies $\mathcal{Y}_1(t, s) > 0$ for all $t \in (s, \infty)_{\mathbb{T}}$ and all $s \in [t_0, \infty)_{\mathbb{T}}$. Rewriting (2.1) in the form

$$\begin{gathered}\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \sum_{i \in [1,n]_{\mathbb{N}}} B_i(t)x(\alpha_i(t)) = f(t) - \sum_{i \in [1,n]_{\mathbb{N}}} [A_i(t) - B_i(t)]x(\alpha_i(t)), \quad t \in [t_0, \infty)_{\mathbb{T}} \\ x(t_0) = x_1, \quad x^{\Delta}(t_0) = x_2, \quad x(t) = \varphi(t), \quad t \in [t_{-1}, t_0)_{\mathbb{T}},\end{gathered} \tag{4.18}$$

applying Lemma 2.4, and using ($A6$), we have

$$\begin{aligned} x(t) &= x_2\mathcal{Y}_1(t, t_0) + x_1\mathcal{Y}_2(t, t_0) + \int_{t_0}^{t} \mathcal{Y}_1(t, \sigma(\eta)) \\ &\quad\times \left[f(\eta) - \sum_{i \in [1,n]_{\mathbb{N}}} [A_i(\eta) - B_i(\eta)]\chi_{[t_0,\infty)_{\mathbb{T}}}(\alpha_i(\eta))x(\alpha_i(\eta)) - \sum_{i \in [1,n]_{\mathbb{N}}} B_i(\eta)\varphi(\alpha_i(\eta))\right]\Delta\eta \\ &\leq y_2\mathcal{Y}_1(t, t_0) + y_1\mathcal{Y}_2(t, t_0) + \int_{t_0}^{t} \mathcal{Y}_1(t, \sigma(\eta))\left[g(\eta) - \sum_{i \in [1,n]_{\mathbb{N}}} B_i(\eta)\psi(\alpha_i(\eta))\right]\Delta\eta \\ &= y(t) \end{aligned} \tag{4.19}$$

for all $t \in [t_0, \infty)_{\mathbb{T}}$. This completes the proof. □

to Contents, page 1

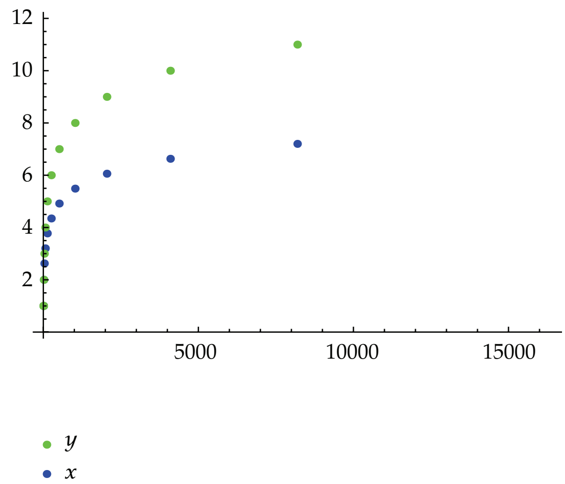


Figure 1: The graph of 10 iterates for the solutions of (4.20) and (4.22) illustrates the result of Theorem 4.4, here $y(t) > x(t)$ for all $t \in (1,\infty)_{\overline{2^{\mathbb{Z}}}}$.

Remark 4.5. If $B_i \in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}}, \mathbb{R}_0^+)$ for $i \in [1,n]_{\mathbb{N}}$, $f(t) \le g(t)$ for all $t \in [t_0,\infty)_{\mathbb{T}}$ and $\varphi(t) \ge \psi(t)$ for all $t \in [t_{-1},t_0)_{\mathbb{T}}$, then $(A6)$ holds.

The following example illustrates Theorem 4.4 for the quantum time scale $\mathbb{T} = \overline{2^{\mathbb{Z}}}$.

Example 4.6. Let $\overline{2^{\mathbb{Z}}} := \{2^k : k \in \mathbb{Z}\} \cup \{0\}$, and consider the following initial value problems:

$$\begin{gathered} D_2\left(\mathrm{Id}_{\overline{2^{\mathbb{Z}}}} D_2 x\right)(t) + \frac{2}{t^4} x\left(\frac{t}{4}\right) = -\frac{1}{t^4} \quad \text{for } t \in [1,\infty)_{\overline{2^{\mathbb{Z}}}}, \\ D_2 x(1) = 1, \quad x(t) \equiv 1 \quad \text{for } t \in \left[\frac{1}{4}, 1\right]_{\overline{2^{\mathbb{Z}}}}, \end{gathered} \tag{4.20}$$

where $\mathrm{Id}_{\overline{2^{\mathbb{Z}}}}$ is the identity function on $\overline{2^{\mathbb{Z}}}$, that is, $\mathrm{Id}_{\overline{2^{\mathbb{Z}}}}(t) = t$ for $t \in \overline{2^{\mathbb{Z}}}$, and

$$D_2 x(t) = \frac{1}{t}\left(x(2t) - x(t)\right) \quad \text{for } t \in \overline{2^{\mathbb{Z}}}, \tag{4.21}$$

$$\begin{gathered} D_2\left(\mathrm{Id}_{\overline{2^{\mathbb{Z}}}} D_2 x\right)(t) + \frac{1}{t^4} x\left(\frac{t}{4}\right) = \frac{1}{t^4} \quad \text{for } t \in [1,\infty)_{\overline{2^{\mathbb{Z}}}}, \\ D_2 x(1) = 1, \quad x(t) \equiv 1 \quad \text{for } t \in \left[\frac{1}{4}, 1\right]_{\overline{2^{\mathbb{Z}}}}. \end{gathered} \tag{4.22}$$

Denoting by x and y the solutions of (4.20) and (4.22), respectively, we obtain $y(t) \ge x(t)$ for all $t \in [1,\infty)_{\overline{2^{\mathbb{Z}}}}$ by Theorem 4.4. For the graph of the first 10 iterates, see Figure 1.

As an immediate consequence of Theorem 4.4, we obtain the following corollary.

to Contents, page 1

Corollary 4.7. *Suppose that (A1), (A2), and (A3) hold and that* (3.1) *is nonoscillatory, then, for* $f \in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}}, \mathbb{R}_0^+)$*, the dynamic equation*

$$\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) x(\alpha_i(t)) = f(t) \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}} \tag{4.23}$$

is also nonoscillatory.

We now consider the following dynamic equation:

$$\begin{gathered}\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) x(\alpha_i(t)) = g(t) \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}, \\ x(t_0) = y_1, \quad x^{\Delta}(t_0) = y_2, \quad x(t) = \psi(t) \quad \text{for } t \in [t_{-1}, t_0)_{\mathbb{T}},\end{gathered} \tag{4.24}$$

where the parameters are the same as in (4.15).

We obtain the most complete result if we compare solutions of (2.1) and (4.24) by omitting the condition $(A2)$ and assuming that the solution of (2.1) is positive.

Corollary 4.8. *Suppose that (A3), (A4), and the following condition hold:*

($A7$) $f, g \in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}}, \mathbb{R})$ *and* $\varphi, \psi \in C_{\mathrm{rd}}([t_{-1}, t_0)_{\mathbb{T}}, \mathbb{R})$ *satisfy*

$$f(t) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\varphi(\alpha_i(t)) \le g(t) - \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\psi(\alpha_i(t)) \quad \forall t \in [t_0,\infty)_{\mathbb{T}}. \tag{4.25}$$

If x is a positive solution of (2.1) *on* $[t_0,\infty)_{\mathbb{T}}$ *with* $x_1 = y_1$ *and* $y_2 \ge x_2$*, then for the solution y of* (4.24)*, one has* $y(t) \ge x(t)$ *for all* $t \in [t_0,\infty)_{\mathbb{T}}$.

Proof. Corollary 4.3 and Remark 3.3 imply that the first fundamental solution $\mathcal{X}_1$ associated with (2.1) (and (4.24)) satisfies $\mathcal{X}_1(t,s) > 0$ for all $t \in (s,\infty)_{\mathbb{T}}$ and all $s \in [t_0,\infty)_{\mathbb{T}}$. Hence, the claim follows from the solution representation formula. □

Remark 4.9. If at least one of the inequalities in the statements of Theorem 4.4 and Corollary 4.8 is strict, then the conclusions hold with the strict inequality too.

Let us compare equations with different coefficients and delays. Now, we consider

$$\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} B_i(t) x(\beta_i(t)) = 0 \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}. \tag{4.26}$$

Theorem 4.10. *Suppose that (A2), (A4), (A5), and the following condition hold:*

($A8$) *for* $i \in [1,n]_{\mathbb{N}}$, $\beta_i \in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}}, \mathbb{T})$ *satisfies* $\beta_i(t) \le \alpha_i(t)$ *for all* $t \in [t_0,\infty)_{\mathbb{T}}$ *and* $\lim_{t\to\infty}\beta_i(t) = \infty$.

Assume further that the first-order dynamic Riccati inequality (3.4) *has a solution* $\Lambda \in C^1_{\mathrm{rd}}([t_1,\infty)_{\mathbb{T}}, \mathbb{R}_0^+)$ *for some* $t_1 \in [t_0,\infty)_{\mathbb{T}}$*, then the first fundamental solution* $\mathcal{Y}_1$ *of* (4.26) *satisfies* $\mathcal{Y}_1(t,s) > 0$ *for all* $t \in (s,\infty)_{\mathbb{T}}$ *and all* $s \in [t_1,\infty)_{\mathbb{T}}$.

to Contents, page 1

Proof. Note that $(A5)$ implies $A_i(t) \geq B_i^+(t)$ for all $t \in [t_0, \infty)_{\mathbb{T}}$ and $i \in [1, n]_{\mathbb{N}}$, then we have

$$\begin{aligned} 0 &\geq \Lambda^{\Delta}(t) + \frac{\Lambda^2(t)}{A_0(t) + \mu(t)\Lambda(t)} + \sum_{i \in [1,n]_{\mathbb{N}}} A_i(t) \mathrm{e}_{\ominus(\Lambda/A_0)}(\sigma(t), \alpha_i(t)) \\ &\geq \Lambda^{\Delta}(t) + \frac{\Lambda^2(t)}{A_0(t) + \mu(t)\Lambda(t)} + \sum_{i \in [1,n]_{\mathbb{N}}} B_i^+(t) \mathrm{e}_{\ominus(\Lambda/A_0)}(\sigma(t), \beta_i(t)) \end{aligned} \tag{4.27}$$

for all $t \in [t_1, \infty)_{\mathbb{T}}$. The reference to Corollary 4.3 (ii) concludes the proof. □

Remark 4.11. If the condition $(A4)$ in Theorem 4.1, Theorem 4.4, Corollary 4.8, and Theorem 4.10 is replaced with $(A1)$, then the claims of the theorems are valid eventually.

Let us introduce the function

$$\alpha_{\max}(t) := \max_{i \in [1,n]_{\mathbb{N}}} \{\alpha_i(t)\} \quad \text{for } t \in [t_0, \infty)_{\mathbb{T}}. \tag{4.28}$$

Corollary 4.12. *Suppose that (A1), (A2), (A3), and (A5) hold. If*

$$\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \left(\sum_{i \in [1,n]_{\mathbb{N}}} A_i(t)\right) x(\alpha_{\max}(t)) = 0 \quad \textit{for } t \in [t_0, \infty)_{\mathbb{T}} \tag{4.29}$$

is nonoscillatory, then (4.1) *is also nonoscillatory.*

Remark 4.13. The claim of Corollary 4.12 is also true when $\alpha_{\max}$ is replaced by σ.

5. Explicit Nonoscillation and Oscillation Results

Theorem 5.1. *Suppose that (A1), (A2), and (A3) hold and that*

$$\frac{\sigma(t)}{2tA_0(t) + \mu(t)} + 2t\sigma(t) \sum_{i \in [1,n]_{\mathbb{N}}} A_i(t) \mathrm{e}_{\ominus(1/(2\mathrm{Id}_{\mathbb{T}} A_0))}(\sigma(t), \alpha_i(t)) \leq 1 \quad \forall t \in [t_1, \infty)_{\mathbb{T}}, \tag{5.1}$$

where $t_1 \in [t_0, \infty)_{\mathbb{T}}$ *and* $\mathrm{Id}_{\mathbb{T}}$ *is the identity function on* $\mathbb{T}$, *then* (3.1) *is nonoscillatory.*

Proof. The statement of the theorem yields that $\Lambda(t) = 1/(2t)$ for $t \in [t_0, \infty)_{\mathbb{T}^+}$ is a positive solution of the Riccati inequality (3.32). □

Next, let us apply Theorem 5.1 to delay differential equations.

to Contents, page 1

Corollary 5.2. *Let $A_0 \in C([t_0,\infty)_{\mathbb{R}},\mathbb{R}^+)$, for $i \in [1,n]_{\mathbb{N}}$, $A_i \in C([t_0,\infty)_{\mathbb{R}},\mathbb{R}_0^+)$, and $\alpha_i \in C([t_0,\infty)_{\mathbb{R}},\mathbb{R})$ such that $\alpha_i(t) \leq t$ for all $t \in [t_0,\infty)_{\mathbb{R}}$ and $\lim_{t\to\infty}\alpha_i(t) = \infty$. If*

$$\frac{1}{2A_0(t)} + 2t^2 \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) \exp\left\{-\int_{\alpha_i(t)}^{t} \frac{1}{2\eta A_0(\eta)} d\eta\right\} \leq 1 \quad \forall t \in [t_1,\infty)_{\mathbb{R}} \tag{5.2}$$

for some $t_1 \in [t_0,\infty)_{\mathbb{R}}$, then (1.2) *is nonoscillatory.*

Now, let us proceed with the discrete case.

Corollary 5.3. *Let $\{A_0(k)\}$ be a positive sequence, for $i \in [1,n]_{\mathbb{N}}$, let $\{A_i(k)\}$ be a nonnegative sequence, and let $\{\alpha_i(k)\}$ be a divergent sequence such that $\alpha_i(k) \leq k+1$ for all $k \in [k_0,\infty)_{\mathbb{N}}$. If*

$$\frac{k+1}{2kA_0(k)+1} + 2k(k+1) \sum_{i\in[1,n]_{\mathbb{N}}} A_i(k) \prod_{j=\alpha_i(k)}^{k} \frac{2jA_0(j)}{2jA_0(j)+1} \leq 1 \quad \forall k \in [k_1,\infty)_{\mathbb{N}} \tag{5.3}$$

for some $k_1 \in [k_0,\infty)_{\mathbb{N}}$, then (1.8) *is nonoscillatory.*

Let us introduce the function

$$A(t,s) := \int_s^t \frac{1}{A_0(\eta)} \Delta\eta \quad \text{for } s,t \in [t_0,\infty)_{\mathbb{T}}. \tag{5.4}$$

Theorem 5.4. *Suppose that (A1), (A2), and (A3) hold, and for every $t_1 \in [t_0,\infty)_{\mathbb{T}}$, the dynamic equation*

$$\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \frac{1}{A(\alpha_{\max}(t),t_1)} \left(\sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) A(\alpha_i(t),t_1)\right) x(\alpha_{\max}(t)) = 0, \quad t \in [t_2,\infty)_{\mathbb{T}} \tag{5.5}$$

is oscillatory, where $t_2 \in [t_1,\infty)_{\mathbb{T}}$ satisfies $\alpha_{\min}(t) > t_1$ for all $t \in [t_2,\infty)_{\mathbb{T}}$, then (3.1) *is also oscillatory.*

Proof. Assume to the contrary that (3.1) is nonoscillatory, then there exists a solution x of (3.1) such that $x > 0$, $(A_0x^{\Delta})^{\Delta} \leq 0$ on $[t_1,\infty)_{\mathbb{T}} \subset [t_0,\infty)_{\mathbb{T}}$. This implies that A_0x^{Δ} is nonincreasing on $[t_1,\infty)_{\mathbb{T}}$, then it follows that

$$x(t) \geq x(t) - x(t_1) = \int_{t_1}^{t} \frac{1}{A_0(\eta)} A_0(\eta)x^{\Delta}(\eta)\Delta\eta \geq A(t,t_1)A_0(t)x^{\Delta}(t) \quad \forall t \in [t_1,\infty)_{\mathbb{T}}, \tag{5.6}$$

or simply by using (5.4),

$$x(t) - A(t,t_1)A_0(t)x^{\Delta}(t) \geq 0 \quad \forall t \in [t_1,\infty)_{\mathbb{T}}. \tag{5.7}$$

Now, let

$$\varphi(t) := \frac{x(t)}{A(t,t_1)} \quad \text{for } t \in (t_1,\infty)_{\mathbb{T}}. \tag{5.8}$$

By the quotient rule, (5.4) and (5.7), we have

$$\varphi^{\Delta}(t) = \frac{A(t,t_1)A_0(t)x^{\Delta}(t) - x(t)}{A(\sigma(t),t_1)A(t,t_1)A_0(t)} \le 0 \quad \forall t \in (t_1,\infty)_{\mathbb{T}}, \tag{5.9}$$

proving that φ is nonincreasing on $(t_1,\infty)_{\mathbb{T}}$. Therefore, for all $i \in [1,n]_{\mathbb{N}}$, we obtain

$$\frac{x(\alpha_{\max}(t))}{A(\alpha_{\max}(t),t_1)} = \varphi(\alpha_{\max}(t)) \le \varphi(\alpha_i(t)) = \frac{x(\alpha_i(t))}{A(\alpha_i(t),t_1)} \quad \forall t \in [t_2,\infty)_{\mathbb{T}}, \tag{5.10}$$

where $t_2 \in [t_1,\infty)_{\mathbb{T}}$ satisfies $\alpha_{\min}(t) > t_1$ for all $t \in [t_2,\infty)_{\mathbb{T}}$. Using (5.10) in (3.1), we see that x solves

$$\left(A_0 x^{\Delta}\right)^{\Delta}(t) + \frac{1}{A(\alpha_{\max}(t),t_1)}\left(\sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)A(\alpha_i(t),t_1)\right)x(\alpha_{\max}(t)) \le 0 \quad \forall t \in [t_2,\infty)_{\mathbb{T}}, \tag{5.11}$$

which shows that (5.5) is also nonoscillatory by Theorem 3.1. This is a contradiction, and the proof is completed. □

The following theorem can be regarded as the dynamic generalization of Leighton's result (Theorem A).

Theorem 5.5. *Suppose that (A2), (A3), and (A4) hold and that*

$$\int_{t_2}^{\infty} \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta)\mathrm{e}_{\ominus(1/(A_0A(\cdot,t_1)))}(\sigma(\eta),\alpha_i(\eta))\Delta\eta = \infty, \tag{5.12}$$

where $t_2 \in (t_1,\infty) \subset [t_0,\infty)_{\mathbb{T}}$, then every solution of (3.1) is oscillatory.

Proof. Assume to the contrary that (3.1) is nonoscillatory. It follows from Theorem 3.1 and Remark 3.2 that (3.4) has a solution $\Lambda \in \mathrm{C_{rd}}([t_0,\infty)_{\mathbb{T}},\mathbb{R}_0^{+})$. Using (3.5) and (5.7), we see that

$$\Lambda(t) \le \frac{1}{A(t,t_1)} \quad \forall t \in [t_2,\infty)_{\mathbb{T}}, \tag{5.13}$$

which together with (3.4) implies that

$$\Lambda^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\mathrm{e}_{\ominus(1/(A_0A(\cdot,t_1)))}(\sigma(t),\alpha_i(t)) \le 0 \quad \forall t \in [t_2,\infty)_{\mathbb{T}}. \tag{5.14}$$

to Contents, page 1

Integrating the last inequality, we get

$$\Lambda(t) - \Lambda(t_2) + \int_{t_2}^{t} \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta) \mathrm{e}_{\ominus(1/(A_0 A(\cdot,t_1)))}(\sigma(\eta), \alpha_i(\eta))\Delta\eta \le 0 \quad \forall t \in [t_2,\infty)_{\mathbb{T}}, \tag{5.15}$$

which is in a contradiction with (5.12). This completes the proof. □

We conclude this section with applications of Theorem 5.5 to delay differential equations and difference equations.

Corollary 5.6. *Let $A_0 \in C([t_0,\infty)_{\mathbb{R}}, \mathbb{R}^+)$, for $i \in [1,n]_{\mathbb{N}}$, $A_i \in C([t_0,\infty)_{\mathbb{R}}, \mathbb{R}_0^+)$, and $\alpha_i \in C([t_0,\infty)_{\mathbb{R}}, \mathbb{R})$ such that $\alpha_i(t) \le t$ for all $t \in [t_0,\infty)_{\mathbb{R}}$ and $\lim_{t\to\infty}\alpha_i(t) = \infty$. If*

$$\lim_{t\to\infty} A(t,t_0) = \infty, \qquad \int_{t_0}^{\infty} \sum_{i\in[1,n]_{\mathbb{N}}} A_i(\eta) \frac{A(\alpha_i(\eta),t_0)}{A(\eta,t_0)} d\eta = \infty, \tag{5.16}$$

where

$$A(t,s) := \int_s^t \frac{1}{A_0(\eta)} d\eta \quad \text{for } s,t \in [t_0,\infty)_{\mathbb{R}}, \tag{5.17}$$

then (1.2) *is oscillatory.*

Corollary 5.7. *Let $\{A_0(k)\}$ be a positive sequence, for $i \in [1,n]_{\mathbb{N}}$, let $\{A_i(k)\}$ be a nonnegative sequence and let $\{\alpha_i(k)\}$ be a divergent sequence such that $\alpha_i(k) \le k+1$ for all $k \in [k_0,\infty)_{\mathbb{N}}$. If*

$$\lim_{k\to\infty} A(k,k_0) = \infty, \qquad \sum_{j=k_0}^{\infty} \sum_{i\in[1,n]_{\mathbb{N}}} A_i(j) \prod_{\ell=\alpha_i(j)}^{j} \frac{A_0(\ell)A(\ell,k_0)}{A_0(\ell)A(\ell,k_0)+1} = \infty, \tag{5.18}$$

where

$$A(k,l) := \sum_{j=l}^{k-1} \frac{1}{A_0(j)} \quad \text{for } l,k \in [k_0,\infty)_{\mathbb{N}}, \tag{5.19}$$

then (1.8) *is oscillatory.*

6. Existence of a Positive Solution

Theorem 6.1. *Suppose that (A2), (A3), and (A4) hold, $f \in C_{\mathrm{rd}}([t_0,\infty)_{\mathbb{T}}, \mathbb{R}_0^+)$, and the first-order dynamic Riccati inequality* (3.4) *has a solution $\Lambda \in C_{\mathrm{rd}}^1([t_0,\infty)_{\mathbb{T}}, \mathbb{R}_0^+)$. Moreover, suppose that there exist $x_1, x_2 \in \mathbb{R}^+$ such that $\varphi(t) \le x_1$ for all $t \in [t_{-1},t_0)_{\mathbb{T}}$ and $x_2 \ge \Lambda(t_0)x_1/A_0(t_0)$, then* (2.1) *admits a positive solution x such that $x(t) \ge x_1$ for all $t \in [t_0,\infty)_{\mathbb{T}}$.*

to Contents, page 1

Proof. First assume that y is the solution of the following initial value problem:

$$\begin{aligned} &\left(A_0 y^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) y(\alpha_i(t)) = 0 \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}, \\ &y^{\Delta}(t_0) = \frac{\Lambda(t_0)}{A_0(t_0)} x_1, \quad y(t) \equiv x_1 \quad \text{for } t \in [t_{-1}, t_0]_{\mathbb{T}}. \end{aligned} \tag{6.1}$$

Denote

$$z(t) := \begin{cases} x_1 \mathrm{e}_{\Lambda/A_0}(t,t_0) & \text{for } t \in [t_0,\infty)_{\mathbb{T}}, \\ x_1 & \text{for } t \in [t_{-1},t_0)_{\mathbb{T}}, \end{cases} \tag{6.2}$$

then, by following similar arguments to those in the proof of the part (ii)⇒(iii) of Theorem 3.1, we obtain

$$\begin{aligned} g(t) &:= \left(A_0 z^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) z(\alpha_i(t)) \\ &= x_1 \mathrm{e}_{\Lambda/A_0}(t,t_0)\left[\Lambda^{\Delta}(t) + \frac{1}{A_0(t)}\Lambda^{\sigma}(t)\Lambda(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t)\mathrm{e}_{\ominus(\Lambda/A_0)}(t,\alpha_i(t))\right] \le 0 \end{aligned} \tag{6.3}$$

for all $t \in [t_0,\infty)_{\mathbb{T}}$. So z is a solution to

$$\begin{aligned} &\left(A_0 z^{\Delta}\right)^{\Delta}(t) + \sum_{i\in[1,n]_{\mathbb{N}}} A_i(t) z(\alpha_i(t)) = g(t) \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}, \\ &z^{\Delta}(t_0) = \frac{\Lambda(t_0)}{A_0(t_0)} x_1, \quad z(t) \equiv x_1 \quad \text{for } t \in [t_{-1}, t_0]_{\mathbb{T}}. \end{aligned} \tag{6.4}$$

Theorem 4.4 implies that $y(t) \ge z(t) \ge x_1 > 0$ for all $t \in [t_0,\infty)_{\mathbb{T}}$. By the hypothesis of the theorem, Theorem 4.4, and Corollary 4.8, we have $x(t) \ge y(t) \ge x_1 > 0$ for all $t \in [t_0,\infty)_{\mathbb{T}}$. This completes the proof for the case $f \equiv 0$ and $g \equiv 0$ on $[t_0,\infty)_{\mathbb{T}}$.

The general case where $f \not\equiv 0$ on $[t_0,\infty)_{\mathbb{T}}$ is also a consequence of Theorem 4.4. □

Let us illustrate the result of Theorem 6.1 with the following example.

Example 6.2. Let $\sqrt{\mathbb{N}_0} := \{\sqrt{k} : k \in \mathbb{N}_0\}$, and consider the following delay dynamic equation:

$$\begin{aligned} &\left(\mathrm{Id}_{\sqrt{\mathbb{N}_0}} x^{\Delta}\right)^{\Delta}(t) + \frac{1}{8t\sqrt{t^2+1}}\left(x(t) + \frac{1}{2}x\left(\sqrt{t^2-1}\right)\right) = \frac{1}{t\sqrt{t^2+1}}, \quad t \in [1,\infty)_{\sqrt{\mathbb{N}_0}}, \\ &x^{\Delta}(1) = 2, \quad x(t) \equiv 2 \quad \text{for } t \in [0,1]_{\sqrt{\mathbb{N}_0}}, \end{aligned} \tag{6.5}$$

to Contents, page 1

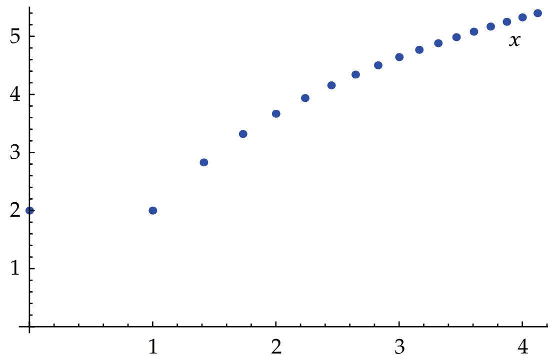


Figure 2: The graph of 15 iterates for the solution of (6.5) illustrates the result of Theorem 6.1.

then (5.1) takes the form $\Phi(t) \leq 1$ for all $t \in [1,\infty)_{\sqrt{\mathbb{N}_0}}$, where the function Φ is defined by

$$\Phi(t) := \frac{1}{2t^2 + \left(\sqrt{t^2+1} - t\right)} \left(\sqrt{t^2+1} + \frac{t^2}{2} \left(1 + \frac{t^2 - 1}{2(t^2-1) + \left(t - \sqrt{t^2-1}\right)} \right) \right) \quad \text{for } t \in [1,\infty)_{\mathbb{R}} \tag{6.6}$$

and is decreasing on $[1,\infty)_{\mathbb{R}}$ and thus is not greater than $\Phi(1) \approx 0.79$, that is, Theorem 5.1 holds. Theorem 6.1 therefore ensures that the solution is positive on $[1,\infty)_{\sqrt{\mathbb{N}_0}}$. For the graph of 15 iterates, see Figure 2.

7. Discussion and Open Problems

We start this section with discussion of explicit nonoscillation conditions for delay differential and difference equations. Let us first consider the continuous case. Corollary 5.6 with $n = 1$ and $\alpha_1(t) = t$ for $t \in [t_0,\infty)_{\mathbb{R}}$ reduces to Theorem A. Nonoscillation part of Kneser's result for (1.4) follows from Corollary 5.2 by letting $n = 1$, $A_0(t) \equiv 1$, and $\alpha_1(t) = t$ for $t \in [t_0,\infty)_{\mathbb{R}}$. Theorem E is obtained by applying Corollary 5.3 to (1.10).

Known nonoscillation tests for difference equations can also be deduced from the results of the present paper. In [18, Lemma 1.2], Chen and Erbe proved that (1.9) is nonoscillatory if and only if there exists a sequence $\{\Lambda(k)\}$ with $A_0(k) + \Lambda(k) > 0$ for all $k \in [k_1,\infty)_{\mathbb{N}}$ and some $k_1 \in [k_0,\infty)_{\mathbb{N}}$ satisfying

$$\Delta\Lambda(k) + \frac{\Lambda^2(k)}{A_0(k) + \Lambda(k)} + A_1(k) \leq 0 \quad \forall k \in [k_1,\infty)_{\mathbb{N}}. \tag{7.1}$$

Since this result is a necessary and sufficient condition, the conclusion of Theorem F could be deduced from

$$\Delta\Lambda(k) + \frac{\Lambda^2(k)}{1 + \Lambda(k)} + A_1(k) \leq 0 \quad \forall k \in [k_1,\infty)_{\mathbb{N}}, \tag{7.2}$$

to Contents, page 1

which is a particular case of (7.1) with $A_0(k) \equiv 1$ for $k \in [k_0, \infty)_{\mathbb{N}}$. We present below a short proof for the nonoscillation part only. Assuming (1.12) and letting

$$\Lambda(k) := \frac{1}{4(k-1)} + \sum_{j=k}^{\infty} A_1(j) \quad \text{for } k \in [k_1, \infty)_{\mathbb{N}} \subset [2, \infty)_{\mathbb{N}}, \tag{7.3}$$

we get

$$\frac{1}{4(k-1)} + \frac{1}{4k} \geq \Lambda(k) \geq \frac{1}{4(k-1)} \quad \forall k \in [k_1, \infty)_{\mathbb{N}}, \tag{7.4}$$

and this yields

$$\Delta\Lambda(k) + \frac{\Lambda^2(k)}{1+\Lambda(k)} + A_1(k) \leq -\frac{1}{4k^2(4k-3)} < 0 \quad \forall\, k \in [k_1, \infty)_{\mathbb{N}}. \tag{7.5}$$

That is, the discrete Riccati inequality (7.2) has a positive solution implying that (1.10) is nonoscillatory. It is not hard to prove that (1.13) implies nonexistence of a sequence $\{\Lambda(k)\}$ satisfying the discrete Riccati inequality (7.2) (see the proof of [23, Lemma 3]). Thus, oscillation/nonoscillation results for (1.10) in [21] can be deduced from nonexistence/existence of a solution for the discrete Riccati inequality (7.2); see also [20].

An application of Theorem 3.1 with $\Lambda(t) := \lambda/t$ for $t \in [t_0, \infty)_{\overline{q^{\mathbb{Z}}}^{+}}$ and $\lambda \in \mathbb{R}^+$ implies the following result for quantum scales.

Example 7.1. Let $\mathbb{T} = \overline{q^{\mathbb{Z}}} := \{q^k : k \in \mathbb{Z}\} \cup \{0\}$ with $q \in (1, \infty)_{\mathbb{R}}$. If there exist $\lambda \in \mathbb{R}_0^+$ and $t_1 \in [t_0, \infty)_{\overline{q^{\mathbb{Z}}}^{+}}$ such that

$$\frac{\lambda^2}{A_0(t) + (q-1)\lambda} + t^2 \sum_{i \in [1,n]_{\mathbb{N}}} A_i(t) \prod_{\eta = \log_q(\alpha_i(t))}^{\log_q(t)} \frac{A_0(q^\eta)}{A_0(q^\eta) + (q-1)\lambda} \leq \frac{\lambda}{q}, \quad t \in [t_1, \infty)_{\overline{q^{\mathbb{Z}}}}, \tag{7.6}$$

then the delay q-difference equation

$$D_q\big(A_0 D_q x\big)(t) + \sum_{i \in [1,n]_{\mathbb{N}}} A_i(t) x(\alpha_i(t)) = 0 \quad \text{for } t \in [t_0, \infty)_{\overline{q^{\mathbb{Z}}}} \tag{7.7}$$

is nonoscillatory.

In [36], Bohner and Ünal studied nonoscillation and oscillation of the q-difference equation

$$D_q^2 x(t) + \frac{a}{qt^2} x(qt) = 0 \quad \text{for } t \in [t_0, \infty)_{\overline{q^{\mathbb{Z}}}}, \tag{7.8}$$

to Contents, page 1

where $a \in \mathbb{R}_0^+$, and proved that (7.7) is nonoscillatory if and only if

$$a \leq \frac{1}{\left(\sqrt{q}+1\right)^2}. \tag{7.9}$$

For the above q-difference equation, (7.6) reduces to the algebraic inequality

$$\frac{\lambda^2}{1+(q-1)\lambda} + \frac{a}{q} \leq \frac{\lambda}{q} \quad \text{or} \quad \lambda^2 - (1-(q-1)a)\lambda + a \leq 0, \tag{7.10}$$

whose discriminant is $(1-(q-1)a)^2 - 4a = (q-1)^2a^2 - (q+1)a + 1$. The discriminant is nonnegative if and only if

$$a \geq \frac{q+2\sqrt{q}+1}{q^2-2q+1} = \frac{1}{\left(\sqrt{q}-1\right)^2} \quad \text{or} \quad a \leq \frac{q-2\sqrt{q}+1}{q^2-2q+1} = \frac{1}{\left(\sqrt{q}+1\right)^2}. \tag{7.11}$$

If the latter one holds, then the inequality (7.6) holds with an equality for the value

$$\lambda := \frac{1}{2}\left(1-(q-1)a + \sqrt{\left(1-(q-1)a\right)^2 - 4a}\right). \tag{7.12}$$

It is easy to check that this value is not less than $2/(\sqrt{q}+1)^2$, that is, the solution is nonnegative. This gives us the nonoscillation part of [36, Theorem 3].

Let us also outline connections to some known results in the theory of second-order ordinary differential equations. For example, the Sturm-Picone comparison theorem is an immediate corollary of Theorem 4.10 if we remark that a solution $\Lambda \in C_{\text{rd}}^1([t_1,\infty)_{\mathbb{T}},\mathbb{R})$ of the inequality (3.32) satisfying $\Lambda/A_0 \in \mathcal{R}^+([t_1,\infty)_{\mathbb{T}},\mathbb{R})$ is also a solution of (3.32) with B_i instead of A_i for $i = 0,1$.

Proposition 7.2 (see [28, 32, 36]). *Suppose that $B_0(t) \geq A_0(t) > 0$, $A_1(t) \geq 0$, and $A_1(t) \geq B_1(t)$ for all $t \in [t_0,\infty)_{\mathbb{T}}$, then nonoscillation of*

$$\left(A_0x^{\Delta}\right)^{\Delta}(t) + A_1(t)x^{\sigma}(t) = 0 \quad \textit{for } t \in [t_0,\infty)_{\mathbb{T}} \tag{7.13}$$

implies nonoscillation of

$$\left(B_0x^{\Delta}\right)^{\Delta}(t) + B_1(t)x^{\sigma}(t) = 0 \quad \textit{for } t \in [t_0,\infty)_{\mathbb{T}}. \tag{7.14}$$

The following result can also be regarded as another generalization of the Sturm-Picone comparison theorem. It is easily deduced that there is a solution $\Lambda \in C_{\text{rd}}^1([t_1,\infty)_{\mathbb{T}},\mathbb{R}_0^+)$ of the inequality (3.4).

to Contents, page 1

Proposition 7.3. *Suppose that (A4) and the conditions of Proposition 7.2 are fulfilled, then nonoscillation of*

$$\left(A_0x^{\Delta}\right)^{\Delta}(t) + A_1(t)x(t) = 0 \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}} \tag{7.15}$$

implies the same for

$$\left(B_0x^{\Delta}\right)^{\Delta}(t) + B_1(t)x(t) = 0 \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}. \tag{7.16}$$

Finally, let us present some open problems. To this end, we will need the following definition.

Definition 7.4. A solution x of (3.1) is said to be *slowly oscillating* if for every $t_1 \in [t_0,\infty)_{\mathbb{T}}$ there exist $t_2 \in (t_1,\infty)_{\mathbb{T}}$ with $\alpha_{\min}(t) \geq t_1$ for all $t \in [t_2,\infty)_{\mathbb{T}}$ and $t_3 \in [t_2,\infty)_{\mathbb{T}}$ such that $x(t_1)x^{\sigma}(t_1) \leq 0$, $x(t_2)x^{\sigma}(t_2) \leq 0$, $x(t) > 0$ for all $t \in (t_1,t_2)_{\mathbb{T}}$.

Following the method of [8, Theorem 10], we can demonstrate that if (A1), (A2) with positive coefficients and (A3) hold, then the existence of a slowly oscillating solution of (3.1) which has infinitely many zeros implies oscillation of all solutions.

(P1) Generally, will existence of a slowly oscillating solution imply oscillation of all solutions? To the best of our knowledge, slowly oscillating solutions have not been studied for difference equations yet, the only known result is [9, Proposition 5.2].

All the results of the present paper are obtained under the assumptions that all coefficients of (3.1) are nonnegative, and if some of them are negative, it is supposed that the equation with the negative terms omitted has a positive solution.

(P2) Obtain sufficient nonoscillation conditions for (3.1) with coefficients of an arbitrary sign, not assuming that all solutions of the equation with negative terms omitted are nonoscillatory. In particular, consider the equation with one oscillatory coefficient.

(P3) Describe the asymptotic and the global properties of nonoscillatory solutions.

(P4) Deduce nonoscillation conditions for linear second-order impulsive equations on time scales, where both the solution and its derivative are subject to the change at impulse points (and these changes can be matched or not). The results of this type for second-order delay differential equations were obtained in [37].

(P5) Consider the same equation on different time scales. In particular, under which conditions will nonoscillation of (1.8) imply nonoscillation of (1.2)?

(P6) Obtain nonoscillation conditions for neutral delay second-order equations. In particular, for difference equations some results of this type (a necessary oscillation conditions) can be found in [17].

(P7) In the present paper, all parameters of the equation are rd-continuous which corresponds to continuous delays and coefficients for differential equations. However, in [8], nonoscillation of second-order equations is studied under a more general assumption that delays and coefficients are Lebesgue measurable functions. Can the restrictions of rd-continuity of the parameters be relaxed to involve,

to Contents, page 1

for example, discontinuous coefficients which arise in the theory of impulsive equations?

Appendix

Time Scales Essentials

A *time scale*, which inherits the standard topology on $\mathbb{R}$, is a nonempty closed subset of reals. Here, and later throughout this paper, a time scale will be denoted by the symbol $\mathbb{T}$, and the intervals with a subscript $\mathbb{T}$ are used to denote the intersection of the usual interval with $\mathbb{T}$. For $t \in \mathbb{T}$, we define the *forward jump operator* $\sigma : \mathbb{T} \to \mathbb{T}$ by $\sigma(t) := \inf(t, \infty)_{\mathbb{T}}$ while the *backward jump operator* $\rho : \mathbb{T} \to \mathbb{T}$ is defined by $\rho(t) := \sup(-\infty, t)_{\mathbb{T}}$, and the *graininess function* $\mu : \mathbb{T} \to \mathbb{R}_0^+$ is defined to be $\mu(t) := \sigma(t) - t$. A point $t \in \mathbb{T}$ is called *right dense* if $\sigma(t) = t$ and/or equivalently $\mu(t) = 0$ holds; otherwise, it is called *right scattered*, and similarly *left dense* and *left scattered* points are defined with respect to the backward jump operator. For $f : \mathbb{T} \to \mathbb{R}$ and $t \in \mathbb{T}$, the Δ-derivative $f^{\Delta}(t)$ of f at the point t is defined to be the number, provided it exists, with the property that, for any $\varepsilon > 0$, there is a neighborhood U of t such that

$$\left| [f^{\sigma}(t) - f(s)] - f^{\Delta}(t)[\sigma(t) - s] \right| \leq \varepsilon |\sigma(t) - s| \quad \forall s \in U, \tag{A.1}$$

where $f^{\sigma} := f \circ \sigma$ on $\mathbb{T}$. We mean the Δ-derivative of a function when we only say derivative unless otherwise is specified. A function f is called *rd-continuous* provided that it is continuous at right-dense points in $\mathbb{T}$ and has a finite limit at left-dense points, and the *set of rd-continuous functions* is denoted by $C_{\mathrm{rd}}(\mathbb{T}, \mathbb{R})$. The set of functions $C^1_{\mathrm{rd}}(\mathbb{T}, \mathbb{R})$ includes the functions whose derivative is in $C_{\mathrm{rd}}(\mathbb{T}, \mathbb{R})$ too. For a function $f \in C^1_{\mathrm{rd}}(\mathbb{T}, \mathbb{R})$, the so-called *simple useful formula* holds

$$f^{\sigma}(t) = f(t) + \mu(t) f^{\Delta}(t) \quad \forall t \in \mathbb{T}^{\kappa}, \tag{A.2}$$

where $\mathbb{T}^{\kappa} := \mathbb{T} \setminus \{\sup \mathbb{T}\}$ if $\sup \mathbb{T} = \max \mathbb{T}$ and satisfies $\rho(\max \mathbb{T}) \neq \max \mathbb{T}$; otherwise, $\mathbb{T}^{\kappa} := \mathbb{T}$. For $s, t \in \mathbb{T}$ and a function $f \in C_{\mathrm{rd}}(\mathbb{T}, \mathbb{R})$, the Δ-integral of f is defined by

$$\int_s^t f(\eta) \Delta\eta = F(t) - F(s) \quad \text{for } s, t \in \mathbb{T}, \tag{A.3}$$

where $F \in C^1_{\mathrm{rd}}(\mathbb{T}, \mathbb{R})$ is an antiderivative of f, that is, $F^{\Delta} = f$ on $\mathbb{T}^{\kappa}$. Table 1 gives the explicit forms of the forward jump, graininess, Δ-derivative, and Δ-integral on the well-known time scales of reals, integers, and the quantum set, respectively.

A function $f \in C_{\mathrm{rd}}(\mathbb{T}, \mathbb{R})$ is called *regressive* if $1 + \mu f \neq 0$ on $\mathbb{T}^{\kappa}$, and *positively regressive* if $1 + \mu f > 0$ on $\mathbb{T}^{\kappa}$. The *set of regressive functions* and the *set of positively regressive functions* are denoted by $\mathcal{R}(\mathbb{T}, \mathbb{R})$ and $\mathcal{R}^+(\mathbb{T}, \mathbb{R})$, respectively, and $\mathcal{R}^-(\mathbb{T}, \mathbb{R})$ is defined similarly.

to Contents, page 1

Table 1: Forward jump, Δ-derivative, and Δ-integral.

$\mathbb{T}$	$\mathbb{R}$	$\mathbb{Z}$	$\overline{q^{\mathbb{Z}}}$, $(q>1)$
$\sigma(t)$	t	$t+1$	qt
$f^{\Delta}(t)$	$f'(t)$	$\Delta f(t)$	$D_q f(t) := (f(qt)-f(t))/((q-1)t)$
$\int_s^t f(\eta)\Delta\eta$	$\int_s^t f(\eta)d\eta$	$\sum_{\eta=s}^{t-1} f(\eta)$	$\int_s^t f(\eta)d_q\eta := (q-1)\sum_{\eta=\log_q(s)}^{\log_q(t/q)} f(q^\eta)q^\eta$

Table 2: The exponential function.

$\mathbb{T}$	$\mathbb{R}$	$\mathbb{Z}$	$\overline{q^{\mathbb{Z}}}$, $(q>1)$
$\mathrm{e}_f(t,s)$	$\exp\{\int_s^t f(\eta)d\eta\}$	$\prod_{\eta=s}^{t-1}(1+f(\eta))$	$\prod_{\eta=\log_q(s)}^{\log_q(t/q)}(1+(q-1)q^\eta f(q^\eta))$

Let $f \in \mathcal{R}(\mathbb{T},\mathbb{R})$, then the *exponential function* $\mathrm{e}_f(\cdot,s)$ on a time scale $\mathbb{T}$ is defined to be the unique solution of the initial value problem

$$\begin{gathered} x^{\Delta}(t) = f(t)x(t) \quad \text{for } t \in \mathbb{T}^{\kappa}, \\ x(s) = 1 \end{gathered} \tag{A.4}$$

for some fixed $s \in \mathbb{T}$. For $h \in \mathbb{R}^+$, set $\mathbb{C}_h := \{z \in \mathbb{C} : z \neq -1/h\}$, $\mathbb{Z}_h := \{z \in \mathbb{C} : -\pi/h < \mathrm{Im}(z) \leq \pi/h\}$, and $\mathbb{C}_0 := \mathbb{Z}_0 := \mathbb{C}$. For $h \in \mathbb{R}_0^+$, we define the *cylinder transformation* $\xi_h : \mathbb{C}_h \to \mathbb{Z}_h$ by

$$\xi_h(z) := \begin{cases} z, & h = 0, \\ \dfrac{1}{h}\mathrm{Log}(1+hz), & h > 0 \end{cases} \tag{A.5}$$

for $z \in \mathbb{C}_h$, then the exponential function can also be written in the form

$$\mathrm{e}_f(t,s) := \exp\left\{\int_s^t \xi_{\mu(\eta)}(f(\eta))\Delta\eta\right\} \quad \text{for } s,t \in \mathbb{T}. \tag{A.6}$$

Table 2 illustrates the explicit forms of the exponential function on some well-known time scales.

The exponential function $\mathrm{e}_f(\cdot,s)$ is strictly positive on $[s,\infty)_{\mathbb{T}}$ if $f \in \mathcal{R}^+([s,\infty)_{\mathbb{T}},\mathbb{R})$, while $\mathrm{e}_f(\cdot,s)$ alternates in sign at right-scattered points of the interval $[s,\infty)_{\mathbb{T}}$ provided that $f \in \mathcal{R}^-([s,\infty)_{\mathbb{T}},\mathbb{R})$. For $h \in \mathbb{R}_0^+$, let $z,w \in \mathbb{C}_h$, the *circle plus* $\oplus_h$ and the *circle minus* $\ominus_h$ are defined by $z \oplus_h w := z + w + hzw$ and $z \ominus_h w := (z-w)/(1+hw)$, respectively. Further throughout the paper, we will abbreviate the operations $\oplus_\mu$ and $\ominus_\mu$ simply by $\oplus$ and $\ominus$, respectively. It is also known that $\mathcal{R}^+(\mathbb{T},\mathbb{R})$ is a subgroup of $\mathcal{R}(\mathbb{T},\mathbb{R})$, that is, $0 \in \mathcal{R}^+(\mathbb{T},\mathbb{R})$, $f,g \in \mathcal{R}^+(\mathbb{T},\mathbb{R})$ implies $f \oplus_\mu g \in \mathcal{R}^+(\mathbb{T},\mathbb{R})$ and $\ominus_\mu f \in \mathcal{R}^+(\mathbb{T},\mathbb{R})$, where $\ominus_\mu f := 0 \ominus_\mu f$ on $\mathbb{T}$.

The readers are referred to [32] for further interesting details in the time scale theory.

to Contents, page 1

Acknowledgment

E. Braverman is partially supported by NSERC Research grant. This work is completed while B. Karpuz is visiting the Department of Statistics and Mathematics, University of Calgary, Canada, in the framework of Doctoral Research Scholarship of the Council of Higher Education of Turkey.

References

[1] R. P. Agarwal, M. Bohner, S. R. Grace, and D. O'Regan, *Discrete Oscillation Theory*, Hindawi Publishing Corporation, New York, NY, USA, 2005.

[2] L. H. Erbe, Q. K. Kong, and B. G. Zhang, *Oscillation Theory for Functional-Differential Equations*, vol. 190 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 1995.

[3] I. Győri and G. Ladas, *Oscillation Theory of Delay Differential Equations with Applications*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, NY, USA, 1991.

[4] G. S. Ladde, V. K. Lakshmikantham, and B. G. Zhang, *Oscillation Theory of Differential Equations with Deviating Arguments*, vol. 110 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 1987.

[5] A. D. Myškis, *Linear Differential Equations with Retarded Argument*, Izdat. "Nauka", Moscow, Russia, 1972.

[6] S. B. Norkin, *Differential Equations of the Second Order with Retarded Argument. Some Problems of the Theory of Vibrations of Systems with Retardation*, vol. 3 of *Translations of Mathematical Monographs*, American Mathematical Society, Providence, RI, USA, 1972.

[7] C. A. Swanson, *Comparison and Oscillation Theory of Linear Differential Equations*, vol. 4 of *Mathematics in Science and Engineering*, Academic Press, New York, NY, USA, 1968.

[8] L. Berezansky and E. Braverman, "Some oscillation problems for a second order linear delay differential equation," *Journal of Mathematical Analysis and Applications*, vol. 220, no. 2, pp. 719–740, 1998.

[9] E. Braverman and B. Karpuz, "Nonoscillation of first-order dynamic equations with several delays," *Advances in Difference Equations*, vol. 2010, Article ID 873459, 22 pages, 2010.

[10] W. Leighton, "On self-adjoint differential equations of second order," *Journal of the London Mathematical Society*, vol. 27, pp. 37–47, 1952.

[11] W. J. Coles, "A simple proof of a well-known oscillation theorem," *Proceedings of the American Mathematical Society*, vol. 19, p. 507, 1968.

[12] A. Wintner, "A criterion of oscillatory stability," *Quarterly of Applied Mathematics*, vol. 7, pp. 115–117, 1949.

[13] A. Kneser, "Untersuchungen über die reellen Nullstellen der Integrale linearer Differentialgleichungen," *Mathematische Annalen*, vol. 42, no. 3, pp. 409–435, 1893.

[14] E. Hille, "Non-oscillation theorems," *Transactions of the American Mathematical Society*, vol. 64, pp. 234–252, 1948.

[15] J. Deng, "Oscillation criteria for second-order linear differential equations," *Journal of Mathematical Analysis and Applications*, vol. 271, no. 1, pp. 283–287, 2002.

[16] C. Huang, "Oscillation and nonoscillation for second order linear differential equations," *Journal of Mathematical Analysis and Applications*, vol. 210, no. 2, pp. 712–723, 1997.

[17] R. N. Rath, J. G. Dix, B. L. S. Barik, and B. Dihudi, "Necessary conditions for the solutions of second order non-linear neutral delay difference equations to be oscillatory or tend to zero," *International Journal of Mathematics and Mathematical Sciences*, vol. 2007, Article ID 60907, 16 pages, 2007.

[18] S. Z. Chen and L. H. Erbe, "Riccati techniques and discrete oscillations," *Journal of Mathematical Analysis and Applications*, vol. 142, no. 2, pp. 468–487, 1989.

[19] S. S. Cheng, T. C. Yan, and H. J. Li, "Oscillation criteria for second order difference equation," *Funkcialaj Ekvacioj*, vol. 34, no. 2, pp. 223–239, 1991.

[20] P. Řehák, "Oscillation and nonoscillation criteria for second order linear difference equations," *Fasciculi Mathematici*, no. 31, pp. 71–89, 2001.

[21] B. G. Zhang and Y. Zhou, "Oscillation and nonoscillation for second-order linear difference equations," *Computers & Mathematics with Applications*, vol. 39, no. 1-2, pp. 1–7, 2000.

to Contents, page 1

[22] Y. Zhou and B. G. Zhang, "Oscillations of delay difference equations in a critical state," *Computers & Mathematics with Applications*, vol. 39, no. 7-8, pp. 71–80, 2000.

[23] X. H. Tang, J. S. Yu, and D. H. Peng, "Oscillation and nonoscillation of neutral difference equations with positive and negative coefficients," *Computers & Mathematics with Applications*, vol. 39, no. 7-8, pp. 169–181, 2000.

[24] B. G. Zhang and S. S. Cheng, "Oscillation criteria and comparison theorems for delay difference equations," *Fasciculi Mathematici*, no. 25, pp. 13–32, 1995.

[25] J. Baoguo, L. Erbe, and A. Peterson, "A Wong-type oscillation theorem for second order linear dynamic equations on time scales," *Journal of Difference Equations and Applications*, vol. 16, no. 1, pp. 15–36, 2010.

[26] O. Došlý and S. Hilger, "A necessary and sufficient condition for oscillation of the Sturm-Liouville dynamic equation on time scales," *Journal of Computational and Applied Mathematics*, vol. 141, no. 1-2, pp. 147–158, 2002.

[27] L. Erbe, "Oscillation criteria for second order linear equations on a time scale," *The Canadian Applied Mathematics Quarterly*, vol. 9, no. 4, p. 345–375 (2002), 2001.

[28] L. Erbe, A. Peterson, and P. Řehák, "Comparison theorems for linear dynamic equations on time scales," *Journal of Mathematical Analysis and Applications*, vol. 275, no. 1, pp. 418–438, 2002.

[29] A. Zafer, "On oscillation and nonoscillation of second-order dynamic equations," *Applied Mathematics Letters*, vol. 22, no. 1, pp. 136–141, 2009.

[30] L. Erbe, A. Peterson, and C. C. Tisdell, "Basic existence, uniqueness and approximation results for positive solutions to nonlinear dynamic equations on time scales," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 69, no. 7, pp. 2303–2317, 2008.

[31] B. Karpuz, "Existence and uniqueness of solutions to systems of delay dynamic equations on time scales," *International Journal of Computer Mathematics*, vol. 10, no. M11, pp. 48–58, 2011.

[32] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales: An Introduction with Applications*, Birkhäuser, Boston, Mass, USA, 2001.

[33] B. Karpuz, "Unbounded oscillation of higher-order nonlinear delay dynamic equations of neutral type with oscillating coefficients," *Electronic Journal of Qualitative Theory of Differential Equations*, vol. 2009, no. 34, pp. 1–14, 2009.

[34] L. Erbe, "Oscillation results for second-order linear equations on a time scale," *Journal of Difference Equations and Applications*, vol. 8, no. 11, pp. 1061–1071, 2002.

[35] J. R. Yan, "Oscillatory properties of second-order differential equations with an "integralwise small" coefficient," *Acta Mathematica Sinica*, vol. 30, no. 2, pp. 206–215, 1987.

[36] M. Bohner and M. Ünal, "Kneser's theorem in q-calculus," *Journal of Physics A*, vol. 38, no. 30, pp. 6729–6739, 2005.

[37] L. Berezansky and E. Braverman, "On oscillation of a second order impulsive linear delay differential equation," *Journal of Mathematical Analysis and Applications*, vol. 233, no. 1, pp. 276–300, 1999.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 393875, 8 pages
doi:10.1155/2011/393875

Research Article

On a Maximal Number of Period Annuli

Yelena Kozmina[1] and Felix Sadyrbaev[2]

[1] *Information Systems and Management Institute, Ludzas Street 91, 1019 Riga, Latvia*
[2] *Institute of Mathematics and Computer Science, University of Latvia Rainis Boulevard 29, 1459 Riga, Latvia*

Correspondence should be addressed to Felix Sadyrbaev, felix@latnet.lv

Received 30 October 2010; Accepted 30 December 2010

Academic Editor: Elena Braverman

We consider equation $x'' + g(x) = 0$, where $g(x)$ is a polynomial, allowing the equation to have multiple period annuli. We detect the maximal number of possible period annuli for polynomials of odd degree and show how the respective optimal polynomials can be constructed.

1. Introduction

Consider equation

$$x'' + g(x) = 0, \tag{1.1}$$

where $g(x)$ is an odd degree polynomial with simple zeros.

The equivalent differential system

$$x' = y, \qquad y' = -g(x) \tag{1.2}$$

has critical points at $(p_i, 0)$, where p_i are zeros of $g(x)$. Recall that a critical point O of (1.2) is *a center* if it has a punctured neighborhood covered with nontrivial cycles.

We will use the following definitions.

Definition 1.1 (see [1]). A *central region* is the largest connected region covered with cycles surrounding O.

to Contents, page 1

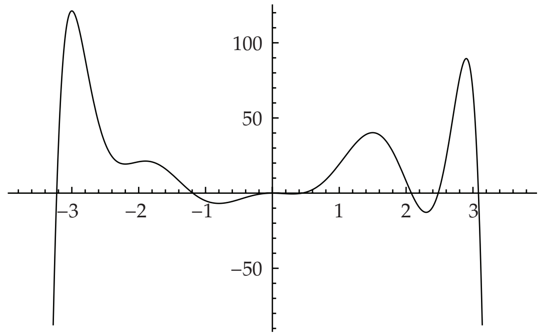


Figure 1

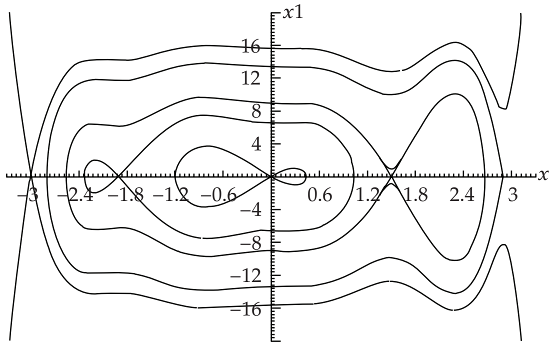


Figure 2: The phase portrait for (1.1), where $G(x)$ is as in Figure 1.

Definition 1.2 (see [1]). A *period annulus* is every connected region covered with nontrivial concentric cycles.

Definition 1.3. We will call a period annulus associated with a central region *a trivial period annulus.* Periodic trajectories of a trivial period annulus encircle exactly one critical point of the type center.

Definition 1.4. Respectively, a period annulus enclosing several (more than one) critical points will be called a *nontrivial period annulus.*

For example, there are four central regions and three nontrivial period annuli in the phase portrait depicted in Figure 2.

Period annuli are the continua of periodic solutions. They can be used for constructing examples of nonlinear equations which have a prescribed number of solutions to the Dirichlet problem

$$x'' + g(x) = 0, \qquad x(0) = 0, \qquad x(1) = 0, \tag{1.3}$$

or a given number of positive solutions [2] to the same problem.

to Contents, page 1

Under certain conditions, period annuli of (1.1) give rise to limit cycles in a dissipative equation

$$x'' + f(x)x' + g(x) = 0. \tag{1.4}$$

The Liénard equation with a quadratical term

$$x'' + f(x)x'^2 + g(x) = 0 \tag{1.5}$$

can be reduced to the form (1.1) by Sabatini's transformation [3]

$$u := \Phi(x) = \int_0^x e^{F(s)} ds, \tag{1.6}$$

where $F(x) = \int_0^x f(s)ds$. Since $du/dx > 0$, this is one-to-one correspondence and the inverse function $x = x(u)$ is well defined.

Lemma 1.5 (see [3, Lemma 1]). *The function $x(t)$ is a solution of* (1.5) *if and only if $u(t) = \Phi(x(t))$ is a solution to*

$$u'' + g(x(u))e^{F(x(u))} = 0. \tag{1.7}$$

Our task in this article is to define the maximal number of nontrivial period annuli for (1.1).

(A) We suppose that $g(x)$ is an odd degree polynomial with simple zeros and with a negative coefficient at the principal term (so $g(-\infty) = +\infty$ and $g(+\infty) = -\infty$). A zero z is called simple if $g(z) = 0$ and $g'(z) \neq 0$.

The graph of a primitive function $G(x) = \int_0^x g(s)ds$ is an even degree polynomial with possible multiple local maxima.

The function $g(x) = -x(x^2 - p^2)(x^2 - q^2)$ is a sample.

We discuss nontrivial period annuli in Section 2. In Section 3, a maximal number of *regular pairs* is detected. Section 4 is devoted to construction of polynomials $g(x)$ which provide the maximal number of *regular pairs* or, equivalently, nontrivial period annuli in (1.1).

2. Nontrivial Period Annuli

The result below provides the criterium for the existence of nontrivial period annuli.

Theorem 2.1 (see [4]). *Suppose that $g(x)$ in* (1.1) *is a polynomial with simple zeros. Assume that M_1 and M_2 ($M_1 < M_2$) are nonneighboring points of maximum of the primitive function $G(x)$. Suppose that any other local maximum of $G(x)$ in the interval (M_1, M_2) is (strictly) less than $\min\{G(M_1); G(M_2)\}$.*

Then, there exists a nontrivial period annulus associated with a pair (M_1, M_2).

to Contents, page 1

It is evident that if $G(x)$ has m pairs of non-neighboring points of maxima then m nontrivial period annuli exist.

Consider, for example, (1.1), where

$$g(x) = -x(x+3)(x+2.2)(x+1.9)(x+0.8)(x-0.3)(x-1.5)(x-2.3)(x-2.9). \tag{2.1}$$

The equivalent system has alternating "saddles" and "centers", and the graph of $G(x)$ is depicted in Figure 1.

There are three pairs of non-neighboring points of maxima and three nontrivial period annuli exist, which are depicted in Figure 2.

3. Polynomials

Consider a polynomial $G(x)$. Points of local maxima x_i and x_j of $G(x)$ are non-neighboring if the interval (x_i, x_j) contains at least one point of local maximum of $G(x)$.

Definition 3.1. Two non-neighboring points of maxima $x_i < x_j$ of $G(x)$ will be called a *regular pair* if $G(x) < \min\{G(x_i), G(x_j)\}$ at any other point of maximum lying in the interval (x_i, x_j).

Theorem 3.2. *Suppose $g(x)$ is a polynomial which satisfies the condition A. Let $G(x)$ be a primitive function for $g(x)$ and n a number of local maxima of $G(x)$.*

Then, the maximal possible number of regular pairs is $n-2$.

Proof. By induction, let $x_1, x_2, \ldots, x_n$ be successive points of maxima of $G(x)$, $x_1 < x_2 < \cdots < x_n$.

(1) Let $n = 3$. The following combinations are possible at three points of maxima:

(a) $G(x_1) \geq G(x_2) \geq G(x_3)$,

(b) $G(x_2) < G(x_1), G(x_2) < G(x_3)$,

(c) $G(x_1) \leq G(x_2) \leq G(x_3)$,

(d) $G(x_2) \geq G(x_1), G(x_2) \geq G(x_3)$.

Only the case (b) provides *a regular pair*. In this case, therefore, the maximal number of *regular pairs* is 1.

(2) Suppose that for any sequence of $n > 3$ ordered points of maxima of $G(x)$ the maximal number of *regular pairs* is $n-2$. Without loss of generality, add to the right one more point of maximum of the function $G(x)$. We get a sequence of $n+1$ consecutive points of maximum $x_1, x_2, \ldots, x_n, x_{n+1}$, $x_1 < x_2 < \ldots < x_n < x_{n+1}$. Let us prove that the maximal number of *regular pairs* is $n-1$. For this, consider the following possible variants.

(a) The couple x_1, x_n is *a regular pair*. If $G(x_1) > G(x_n)$ and $G(x_{n+1}) > G(x_n)$, then, beside the *regular pairs* in the interval $[x_1, x_n]$, only one new *regular pair* can appear, namely, x_1, x_{n+1}. Then, the maximal number of *regular pairs* which can be composed of the points $x_1, x_2, \ldots, x_n, x_{n+1}$, is not greater than $(n-2)+1 = n-1$. If $G(x_1) \leq G(x_n)$ or $G(x_{n+1}) \leq G(x_n)$, then the additional *regular pair* does not appear. In a particular case $G(x_2) < G(x_3) < \cdots < G(x_n) < G(x_{n+1})$ and $G(x_1) > G(x_n)$ the following *regular pairs* exist, namely, x_1 and x_3, x_1 and $x_4, \ldots$, x_1 and x_n, and the new pair x_1 and x_{n+1} appears, totally $n-1$ pairs.

to Contents, page 1

(b) Suppose that x_1, x_n is not *a regular pair*. Let x_i and x_j be a *regular pair*, $1 \le i < j \le n$, and there is no other *regular pair* x_p, x_q such that $1 \le p \le i < j \le q \le n$. Let us mention that if such a pair x_i, x_j does not exist, then the function $G(x)$ does not have *regular pairs* at all and the sequence $\{G(x_k)\}$, $k = 1, \cdots, n$, is monotone. Then, if $G(x_{n+1})$ is greater than any other maximum, there are exactly $(n+1) - 2 = n - 1$ *regular pairs*.

Otherwise, we have two possibilities:

$$\text{either } G(x_i) \ge G(x_p), \quad p = 1, \ldots, i-1,$$
$$\text{or } G(x_j) \ge G(x_q), \quad q = j+1, \ldots, n.$$

In the first case, the interval $[x_1, x_i]$ contains i points of maximum of $G(x)$, $i < n$, and hence the number of *regular pairs* in this interval does not exceed $i - 2$. There are no *regular pairs* x_p, x_k for $1 \le p < i$, $i < k \le n+1$. The interval $[x_i, x_{n+1}]$ contains $(n+1) - (i-1)$ points of maximum of $G(x)$, and hence the number of *regular pairs* in this interval does not exceed $(n+1) - (i-1) - 2 = n - i$. Totally, there are no more *regular pairs* than $(i-2) + (n-i) = n-2$.

In the second case, the number of *regular pairs* in $[x_i, x_j]$ does not exceed $j - (i-1) - 2 = j - i - 1$. In $[x_j, x_{n+1}]$,there are no more than $(n+1) - (j-1) - 2 = n - j$ *regular pairs*. The points x_p, $p = 1, \ldots, i-1$, x_q, $j < q \le n$ do not form *regular pairs*, by the choice of x_p and x_q. The points x_p, $p = 1, \ldots, i$, together with x_{n+1} (it serves as the $i+1$th point in a collection of points) form not more than $(i+1) - 2 = i - 1$ *regular pairs.* Totally, the number of *regular pairs* is not greater than $(j-i-1) + (n-j) + (i-1) = n-2$. □

4. Existence of Polynomials with Optimal Distribution

Theorem 4.1. *Given number n, a polynomial $g(x)$ can be constructed such that*

(a) *the condition (A) is satisfied,*

(b) *the primitive function $G(x)$ has exactly n points of maximum and the number of regular pairs is exactly $n-2$.*

Proof of the Theorem. Consider the polynomial

$$G(x) = -\left(x+\frac{1}{2}\right)\left(x-\frac{1}{2}\right)\left(x+\frac{3}{2}\right)\left(x-\frac{3}{2}\right)\left(x+\frac{5}{2}\right)\left(x-\frac{5}{2}\right)\left(x+\frac{7}{2}\right)\left(x-\frac{7}{2}\right). \tag{4.1}$$

It is an even function with the graph depicted in Figure 3.

Consider now the polynomial

$$G_\varepsilon(x) = -\left(x+\frac{1}{2}+\varepsilon\right)\left(x-\frac{1}{2}\right)\left(x+\frac{3}{2}\right)\left(x-\frac{3}{2}\right)\left(x+\frac{5}{2}\right)\left(x-\frac{5}{2}\right)\left(x+\frac{7}{2}\right)\left(x-\frac{7}{2}\right), \tag{4.2}$$

where $\varepsilon > 0$ is small enough. The graph of $G_\varepsilon(x)$ with $\varepsilon = 0.2$ is depicted in Figure 4.

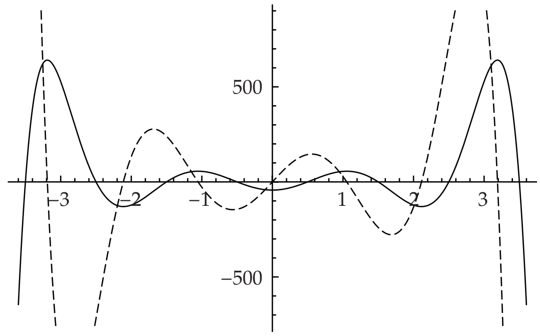


Figure 3: $G(x)$ (solid) and $G'(x) = g(x)$ (dashed).

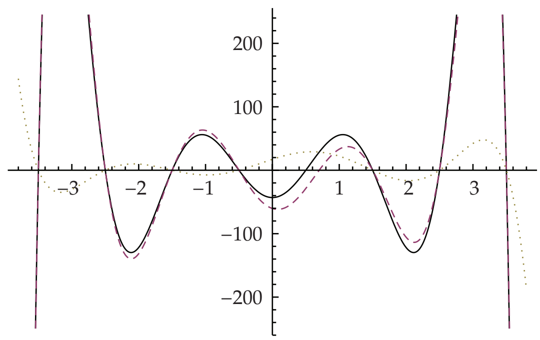


Figure 4: $G(x)$ (solid line), $G_\varepsilon(x)$ (dashed line), and $G(x) - G_\varepsilon(x)$ (dotted line).

Denote the maximal values of $G(x)$ and $G_\varepsilon(x)$ to the right of $x = 0$ m_1^+, m_2^+. Denote the maximal values of $G(x)$ and $G_\varepsilon(x)$ to the left of $x = 0$ m_1^-, m_2^-. One has for $G(x)$ that $m_1^+ = m_1^- < m_2^- = m_2^+$. One has for $G_\varepsilon(x)$ that $m_1^+ < m_1^- < m_2^+ < m_2^-$. Then, there are two regular pairs (resp., m_1^- and m_2^+, m_2^+ and m_2^-).

For arbitrary even n the polynomial

$$G_\varepsilon(x) = -\left(x + \frac{1}{2}\right)\left(x - \frac{1}{2}\right)\left(x + \frac{3}{2}\right)\left(x - \frac{3}{2}\right)\cdots\left(x + \frac{2n-1}{2}\right)\left(x - \frac{2n-1}{2}\right), \tag{4.3}$$

is to be considered where the maximal values $m_1^+, m_2^+, \ldots, m_{n/2}^+$ to the right of $x = 0$ form ascending sequence, and, respectively, the maximal values $m_1^-, m_2^-, \ldots, m_{n/2}^-$ to the left of $x = 0$ also form ascending sequence. One has that $m_i^+ = m_i^-$ for all i. For a slightly modified polynomial

$$G_\varepsilon(x) = -\left(x + \frac{1}{2} + \varepsilon\right)\left(x - \frac{1}{2}\right)\left(x + \frac{3}{2}\right)\left(x - \frac{3}{2}\right)\ldots\left(x + \frac{2n-1}{2}\right)\left(x - \frac{2n-1}{2}\right), \tag{4.4}$$

the maximal values are arranged as

$$m_1^+ < m_1^- < m_2^+ < m_2^- < \cdots < m_{n/2}^+ < m_{n/2}^-. \tag{4.5}$$

to Contents, page 1

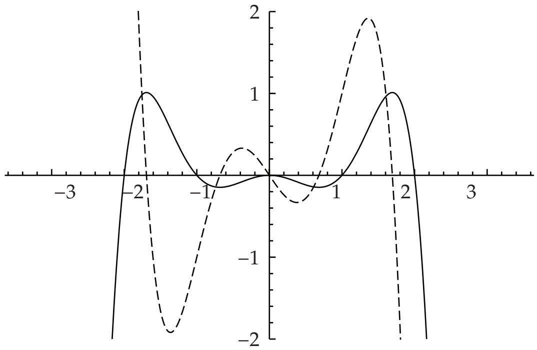


Figure 5: $G(x)$ (solid) and $G'(x) = g(x)$ (dashed).

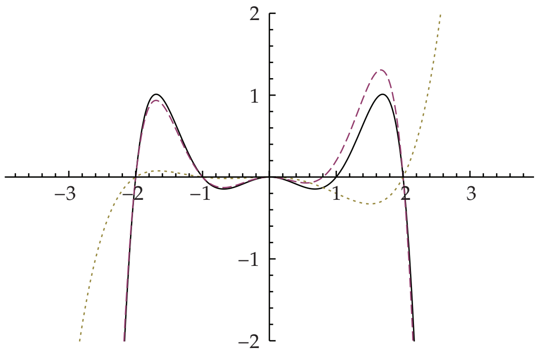


Figure 6: $G(x)$ (solid), $G_\varepsilon(x)$ (dashed), and $G(x) - G_\varepsilon(x)$ (dotted).

Therefore, there exist exactly $n - 2$ *regular pairs* and, consequently, $n - 2$ nontrivial period annuli in the differential equation (1.1).

If n is odd, then the polynomial

$$G(x) = -x^2(x-1)(x+1)(x-2)(x+2)\cdots(x-(n-1))(x+(n-1)) \tag{4.6}$$

with n local maxima is to be considered. The maxima are descending for $x < 0$ and ascending if $x > 0$. The polynomial with three local maxima is depicted in Figure 5.

The slightly modified polynomial

$$G(x) = -x^2(x-1-\varepsilon)(x+1)(x-2)(x+2)\cdots(x-(n-1))(x+(n-1)) \tag{4.7}$$

has maxima which are not equal and are arranged in an optimal way in order to produce the maximal $(n-2)$ *regular pairs*.

The graph of $G_\varepsilon(x)$ with $\varepsilon = 0.2$ is depicted in Figure 6. □

to Contents, page 1

Acknowledgments

This work has been supported by ERAF project no. 2010/0206/2DP/2.1.1.2.0/10/APIA/VIAA/011 and Latvian Council of Science Grant no. 09.1220.

References

[1] M. Sabatini, "Liénard limit cycles enclosing period annuli, or enclosed by period annuli," *The Rocky Mountain Journal of Mathematics*, vol. 35, no. 1, pp. 253–266, 2005.

[2] S. Atslega and F. Sadyrbaev, "Period annuli and positive solutions of nonlinear boundary value problems," in *Proceedings of the 7th Congress of The International Society for Analysis, Its Applications and Computation (ISAAC '09)*, July 2009, http://www.isaac2009.org/Congress/Welcome.html.

[3] M. Sabatini, "On the period function of $x'' + f(x)x'^2 + g(x) = 0$," *Journal of Differential Equations*, vol. 196, no. 1, pp. 151–168, 2004.

[4] S. Atslega and F. Sadyrbaev, "Multiple solutions of the second order nonlinear Neumann BVP," in *Proceedings of the 6th International Conference on Differential Equations and Dynamical Systems*, pp. 100–103, Watam Press, Baltimore, Md, USA, May 2009.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 637142, 14 pages
doi:10.1155/2011/637142

Research Article

On Nonoscillation of Advanced Differential Equations with Several Terms

L. Berezansky[1] and E. Braverman[2]

[1] *Department of Mathematics, Ben-Gurion University of Negev, Beer-Sheva 84105, Israel*
[2] *Department of Mathematics and Statistics, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada T2N 1N4*

Correspondence should be addressed to E. Braverman, maelena@math.ucalgary.ca

Received 16 December 2010; Accepted 27 January 2011

Academic Editor: Josef Diblík

Existence of positive solutions for advanced equations with several terms $\dot{x}(t) + \sum_{k=1}^{m} a_k(t)x(h_k(t)) = 0$, $h_k(t) \geq t$ is investigated in the following three cases: (a) all coefficients a_k are positive; (b) all coefficients a_k are negative; (c) there is an equal number of positive and negative coefficients. Results on asymptotics of nonoscillatory solutions are also presented.

1. Introduction

This paper deals with nonoscillation properties of scalar advanced differential equations. Advanced differential equations appear in several applications, especially as mathematical models in economics; an advanced term may, for example, reflect the dependency on anticipated capital stock [1, 2].

It is not quite clear how to formulate an initial value problem for such equations, and existence and uniqueness of solutions becomes a complicated issue. To study oscillation, we need to assume that there exists a solution of such equation on the halfline. In the beginning of 1980s, sufficient oscillation conditions for first-order linear advanced equations with constant coefficients and deviations of arguments were obtained in [3] and for nonlinear equations in [4]. Later oscillation properties were studied for other advanced and mixed differential equations (see the monograph [5], the papers [6–12] and references therein). Overall, these publications mostly deal with sufficient oscillation conditions; there are only few results [7, 9, 12] on existence of positive solutions for equations with several advanced terms and variable coefficients, and the general nonoscillation theory is not complete even for first-order linear equations with variable advanced arguments and variable coefficients of the same sign.

to Contents, page 1

The present paper partially fills up this gap. We obtain several nonoscillation results for advanced equations using the generalized characteristic inequality [13]. The main method of this paper is based on fixed point theory; thus, we also state the existence of a solution in certain cases.

In the linear case, the best studied models with advanced arguments were the equations of the types

$$\dot{x}(t) - a(t)x(h(t)) + b(t)x(t) = 0,$$
$$\dot{x}(t) - a(t)x(t) + b(t)x(g(t)) = 0, \tag{1.1}$$

where $a(t) \geq 0$, $b(t) \geq 0$, $h(t) \geq t$, and $g(t) \geq t$.

Let us note that oscillation of higher order linear and nonlinear equations with advanced and mixed arguments was also extensively investigated, starting with [14]; see also the recent papers [15–19] and references therein.

For equations with an advanced argument, the results obtained in [20, 21] can be reformulated as Theorems A–C below.

Theorem A (see [20]). *If a, b, and h are equicontinuous on $[0, \infty)$, $a(t) \geq 0$, $b(t) \geq 0$, $h(t) \geq t$, and $\limsup_{t\to\infty}[h(t) - t] < \infty$, then the advanced equation*

$$\dot{x}(t) + a(t)x(h(t)) + b(t)x(t) = 0 \tag{1.2}$$

has a nonoscillatory solution.

In the present paper, we extend Theorem A to the case of several deviating arguments and coefficients (Theorem 2.10).

Theorem B (see [20]). *If a, b, and h are equicontinuous on $[0, \infty)$, $a(t) \geq 0$, $b(t) \geq 0$, $h(t) \geq t$, $\limsup_{t\to\infty}[h(t) - t] < \infty$, and*

$$\limsup_{t\to\infty} \int_t^{h(t)} a(s) \exp\left\{\int_s^{h(s)} b(\tau)d\tau\right\} ds < \frac{1}{e}, \tag{1.3}$$

then the advanced equation

$$\dot{x}(t) - a(t)x(h(t)) - b(t)x(t) = 0 \tag{1.4}$$

has a nonoscillatory solution.

Corollary 2.3 of the present paper extends Theorem B to the case of several coefficients $a_k \geq 0$ and advanced arguments h_k (generally, $b(t) \equiv 0$); if

$$\int_t^{\max_k h_k(t)} \sum_{i=1}^{m} a_i(s)ds \leq \frac{1}{e}, \tag{1.5}$$

to Contents, page 1

then the equation

$$\dot{x}(t) + \sum_{k=1}^{m} a_k(t)x(h_k(t)) = 0 \tag{1.6}$$

has an eventually positive solution. To the best of our knowledge, only the opposite inequality (with $\min_k h_k(t)$ rather than $\max_k h_k(t)$ in the upper bound) was known as a sufficient oscillation condition. Coefficients and advanced arguments are also assumed to be of a more general type than in [20]. Comparison to equations with constant arguments deviations, and coefficients (Corollary 2.8) is also outlined.

For advanced equations with coefficients of different sign, the following result is known.

Theorem C (see [21]). *If $0 \le a(t) \le b(t)$ and $h(t) \ge t$, then there exists a nonoscillatory solution of the equation*

$$\dot{x}(t) - a(t)x(h(t)) + b(t)x(t) = 0. \tag{1.7}$$

This result is generalized in Theorem 2.13 to the case of several positive and negative terms and several advanced arguments; moreover, positive terms can also be advanced as far as the advance is not greater than in the corresponding negative terms.

We also study advanced equations with positive and negative coefficients in the case when positive terms dominate rather than negative ones; some sufficient nonoscillation conditions are presented in Theorem 2.15; these results are later applied to the equation with constant advances and coefficients. Let us note that analysis of nonoscillation properties of the mixed equation with a positive advanced term

$$\dot{x}(t) + a(t)x(h(t)) - b(t)x\big(g(t)\big) = 0, \quad h(t) \ge t,\ g(t) \le t,\ a(t) \ge 0,\ b(t) \ge 0 \tag{1.8}$$

was also more complicated compared to other cases of mixed equations with positive and negative coefficients [21].

In nonoscillation theory, results on asymptotic properties of nonoscillatory solutions are rather important; for example, for equations with several delays and positive coefficients, all nonoscillatory solutions tend to zero if the integral of the sum of coefficients diverges; under the same condition for negative coefficients, all solutions tend to infinity. In Theorems 2.6 and 2.11, the asymptotic properties of nonoscillatory solutions for advanced equations with coefficients of the same sign are studied.

The paper is organized as follows. Section 2 contains main results on the existence of nonoscillatory solutions to advanced equations and on asymptotics of these solutions: first for equations with coefficients of the same sign, then for equations with both positive and negative coefficients. Section 3 involves some comments and open problems.

to Contents, page 1

2. Main Results

Consider first the equation

$$\dot{x}(t) - \sum_{k=1}^{m} a_k(t)x(h_k(t)) = 0, \tag{2.1}$$

under the following conditions:

(a1) $a_k(t) \geq 0$, $k = 1, \ldots, m$, are Lebesgue measurable functions locally essentially bounded for $t \geq 0$,

(a2) $h_k : [0, \infty) \to \mathbb{R}$ are Lebesgue measurable functions, $h_k(t) \geq t$, $k = 1, \ldots, m$.

Definition 2.1. A locally absolutely continuous function $x : [t_0, \infty) \to R$ is called *a solution of problem* (2.1) if it satisfies (2.1) for almost all $t \in [t_0, \infty)$.

The same definition will be used for all further advanced equations.

Theorem 2.2. *Suppose that the inequality*

$$u(t) \geq \sum_{k=1}^{m} a_k(t) \exp\left\{ \int_t^{h_k(t)} u(s)ds \right\}, \quad t \geq t_0 \tag{2.2}$$

has a nonnegative solution which is integrable on each interval $[t_0, b]$, *then* (2.1) *has a positive solution for* $t \geq t_0$.

Proof. Let $u_0(t)$ be a nonnegative solution of inequality (2.2). Denote

$$u_{n+1}(t) = \sum_{k=1}^{m} a_k(t) \exp\left\{ \int_t^{h_k(t)} u_n(s)ds \right\}, \quad n = 0, 1, \ldots, \tag{2.3}$$

then

$$u_1(t) = \sum_{k=1}^{m} a_k(t) \exp\left\{ \int_t^{h_k(t)} u_0(s)ds \right\} \leq u_0(t). \tag{2.4}$$

By induction we have $0 \leq u_{n+1}(t) \leq u_n(t) \leq u_0(t)$. Hence, there exists a pointwise limit $u(t) = \lim_{n\to\infty} u_n(t)$. By the Lebesgue convergence theorem, we have

$$u(t) = \sum_{k=1}^{m} a_k(t) \exp\left\{ \int_t^{h_k(t)} u(s)ds \right\}. \tag{2.5}$$

Then, the function

$$x(t) = x(t_0) \exp\left\{ \int_{t_0}^{t} u(s)ds \right\} \quad \text{for any } x(t_0) > 0 \tag{2.6}$$

is a positive solution of (2.1). □

Corollary 2.3. *If*

$$\int_t^{\max_k h_k(t)} \sum_{i=1}^m a_i(s)ds \le \frac{1}{e}, \quad t \ge t_0, \tag{2.7}$$

then (2.1) *has a positive solution for* $t \ge t_0$.

Proof. Let $u_0(t) = e\sum_{k=1}^m a_k(t)$, then u_0 satisfies (2.2) at any point t where $\sum_{k=1}^m a_k(t) = 0$. In the case when $\sum_{k=1}^m a_k(t) \ne 0$, inequality (2.7) implies

$$\begin{aligned}
&\frac{u_0(t)}{\sum_{k=1}^m a_k(t)\exp\left\{\int_t^{h_k(t)} u_0(s)ds\right\}} \\
&\quad \ge \frac{u_0(t)}{\sum_{k=1}^m a_k(t)\exp\left\{\int_t^{\max_k h_k(t)} u_0(s)ds\right\}} \\
&\quad = \frac{e\sum_{k=1}^m a_k(t)}{\sum_{k=1}^m a_k(t)\exp\left\{e\int_t^{\max_k h_k(t)} \sum_{i=1}^m a_i(s)ds\right\}} \\
&\quad \ge \frac{e\sum_{k=1}^m a_k(t)}{\sum_{k=1}^m a_k(t)e} = 1.
\end{aligned} \tag{2.8}$$

Hence, $u_0(t)$ is a positive solution of inequality (2.2). By Theorem 2.2, (2.1) has a positive solution for $t \ge t_0$. □

Corollary 2.4. *If there exists* $\sigma > 0$ *such that* $h_k(t) - t \le \sigma$ *and* $\int_0^\infty \sum_{k=1}^m a_k(s)ds < \infty$, *then* (2.1) *has an eventually positive solution.*

Corollary 2.5. *If there exists* $\sigma > 0$ *such that* $h_k(t) - t \le \sigma$ *and* $\lim_{t\to\infty} a_k(t) = 0$, *then* (2.1) *has an eventually positive solution.*

Proof. Under the conditions of either Corollary 2.4 or Corollary 2.5, obviously there exists $t_0 \ge 0$ such that (2.7) is satisfied. □

Theorem 2.6. *Let* $\int^\infty \sum_{k=1}^m a_k(s)ds = \infty$ *and* x *be an eventually positive solution of* (2.1), *then* $\lim_{t\to\infty} x(t) = \infty$.

Proof. Suppose that $x(t) > 0$ for $t \ge t_1$, then $\dot{x}(t) \ge 0$ for $t \ge t_1$ and

$$\dot{x}(t) \ge \sum_{k=1}^m a_k(t)x(t_1), \quad t \ge t_1, \tag{2.9}$$

which implies

$$x(t) \ge x(t_1)\int_{t_1}^t \sum_{k=1}^m a_k(s)ds. \tag{2.10}$$

Thus, $\lim_{t\to\infty} x(t) = \infty$. □

to Contents, page 1

Consider together with (2.1) the following equation:

$$\dot{x}(t) - \sum_{k=1}^{m} b_k(t) x\big(g_k(t)\big) = 0, \tag{2.11}$$

for $t \geq t_0$. We assume that for (2.11) conditions (a1)-(a2) also hold.

Theorem 2.7. *Suppose that $t \leq g_k(t) \leq h_k(t)$, $0 \leq b_k(t) \leq a_k(t)$, $t \geq t_0$, and the conditions of Theorem 2.2 hold, then* (2.11) *has a positive solution for $t \geq t_0$.*

Proof. Let $u_0(t) \geq 0$ be a solution of inequality (2.2) for $t \geq t_0$, then this function is also a solution of this inequality if $a_k(t)$ and $h_k(t)$ are replaced by $b_k(t)$ and $g_k(t)$. The reference to Theorem 2.2 completes the proof. □

Corollary 2.8. *Suppose that there exist $a_k > 0$ and $\sigma_k > 0$ such that $0 \leq a_k(t) \leq a_k$, $t \leq h_k(t) \leq t + \sigma_k$, $t \geq t_0$, and the inequality*

$$\lambda \geq \sum_{k=1}^{m} a_k e^{\lambda \sigma_k} \tag{2.12}$$

has a solution $\lambda \geq 0$, then (2.1) *has a positive solution for $t \geq t_0$.*

Proof. Consider the equation with constant parameters

$$\dot{x}(t) - \sum_{k=1}^{m} a_k x(t + \sigma_k) = 0. \tag{2.13}$$

Since the function $u(t) \equiv \lambda$ is a solution of inequality (2.2) corresponding to (2.13), by Theorem 2.2, (2.13) has a positive solution. Theorem 2.7 implies this corollary. □

Corollary 2.9. *Suppose that $0 \leq a_k(t) \leq a_k$, $t \leq h_k(t) \leq t + \sigma$ for $t \geq t_0$, and*

$$\sum_{k=1}^{m} a_k \leq \frac{1}{e\sigma}, \tag{2.14}$$

then (2.1) *has a positive solution for $t \geq t_0$.*

Proof. Since $\sum_{k=1}^{m} a_k \leq 1/e\sigma$, the number $\lambda = 1/\sigma$ is a positive solution of the inequality

$$\lambda \geq \left(\sum_{k=1}^{m} a_k\right) e^{\lambda \sigma}, \tag{2.15}$$

which completes the proof. □

to Contents, page 1

Consider now the equation with positive coefficients

$$\dot{x}(t) + \sum_{k=1}^{m} a_k(t) x(h_k(t)) = 0. \tag{2.16}$$

Theorem 2.10. *Suppose that $a_k(t) \geq 0$ are continuous functions bounded on $[t_0, \infty)$ and h_k are equicontinuous functions on $[t_0, \infty)$ satisfying $0 \leq h_k(t) - t \leq \delta$, then* (2.16) *has a nonoscillatory solution.*

Proof. In the space $C[t_0, \infty)$ of continuous functions on $[t_0, \infty)$, consider the set

$$M = \left\{ u \mid 0 \leq u \leq \sum_{k=1}^{m} a_k(t) \right\}, \tag{2.17}$$

and the operator

$$(Hu)(t) = \sum_{k=1}^{m} a_k(t) \exp\left\{ -\int_t^{h_k(t)} u(s) ds \right\}. \tag{2.18}$$

If $u \in M$, then $Hu \in M$.

For the integral operator

$$(Tu)(t) := \int_t^{h_k(t)} u(s) ds, \tag{2.19}$$

we will demonstrate that TM is a compact set in the space $C[t_0, \infty)$. If $u \in M$, then

$$\|(Tu)(t)\|_{C[t_0,\infty)} \leq \sup_{t \geq t_0} \int_t^{t+\delta} |u(s)| ds \leq \sup_{t \geq t_0} \sum_{k=1}^{m} a_k(t) \delta < \infty. \tag{2.20}$$

Hence, the functions in the set TM are uniformly bounded in the space $C[t_0, \infty)$.

Functions h_k are equicontinuous on $[t_0, \infty)$, so for any $\varepsilon > 0$, there exists a $\sigma_0 > 0$ such that for $|t - s| < \sigma_0$, the inequality

$$|h_k(t) - h_k(s)| < \frac{\varepsilon}{2} \left(\sup_{t \geq t_0} \sum_{k=1}^{m} a_k(t) \right)^{-1}, \quad k = 1, \ldots, m \tag{2.21}$$

holds. From the relation

$$\int_{t_0}^{h_k(t_0)} - \int_t^{h_k(t)} = \int_{t_0}^{t} + \int_t^{h_k(t_0)} - \int_t^{h_k(t_0)} - \int_{h_k(t_0)}^{h_k(t)} = \int_{t_0}^{t} - \int_{h_k(t_0)}^{h_k(t)}, \tag{2.22}$$

to Contents, page 1

we have for $|t-t_0| < \min\{\sigma_0, \varepsilon/2\sup_{t\geq t_0}\sum_{k=1}^m a_k(t)\}$ and $u\in M$ the estimate

$$\begin{aligned}
|(Tu)(t)-(Tu)(t_0)| &= \left|\int_{\mathrm{t}}^{h_k(t)} u(s) - \int_{t_0}^{h_k(t_0)} u(s)ds\right| \\
&\leq \int_{t_0}^{t} |u(s)|ds + \int_{h_k(t_0)}^{h_k(t)} |u(s)|ds \\
&\leq |t-t_0| \sup_{t\geq t_0}\sum_{k=1}^{m} a_k(t) + |h_k(t)-h_k(t_0)| \sup_{t\geq t_0}\sum_{k=1}^{m} a_k(t) \\
&< \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon.
\end{aligned} \tag{2.23}$$

Hence, the set TM contains functions which are uniformly bounded and equicontinuous on $[t_0,\infty)$, so it is compact in the space $C[t_0,\infty)$; thus, the set HM is also compact in $C[t_0,\infty)$.

By the Schauder fixed point theorem, there exists a continuous function $u\in M$ such that $u = Hu$, then the function

$$x(t) = \exp\left\{-\int_{t_0}^{t} u(s)ds\right\} \tag{2.24}$$

is a bounded positive solution of (2.16). Moreover, since u is nonnegative, this solution is nonincreasing on $[t_0,\infty)$. □

Theorem 2.11. *Suppose that the conditions of Theorem 2.10 hold,*

$$\int_{t_0}^{\infty}\sum_{k=1}^{m} a_k(s)ds = \infty, \tag{2.25}$$

and x is a nonoscillatory solution of (2.16)*, then* $\lim_{t\to\infty} x(t) = 0$.

Proof. Let $x(t) > 0$ for $t\geq t_0$, then $\dot{x}(t)\leq 0$ for $t\geq t_0$. Hence, $x(t)$ is nonincreasing and thus has a finite limit. If $\lim_{t\to\infty}x(t) = d > 0$, then $x(t) > d$ for any t, and thus $\dot{x}(t) \leq -d\sum_{k=1}^m a_k(t)$ which implies $\lim_{t\to\infty}x(t) = -\infty$. This contradicts to the assumption that $x(t)$ is positive, and therefore $\lim_{t\to\infty}x(t) = 0$. □

Let us note that we cannot guarantee any (exponential or polynomial) rate of convergence to zero even for constant coefficients a_k, as the following example demonstrates.

Example 2.12. Consider the equation $\dot{x}(t) + x(h(t)) = 0$, where $h(t) = t^{t\ln t}$, $t\geq 3$, $x(3) = 1/\ln 3$. Then, $x(t) = 1/(\ln t)$ is the solution which tends to zero slower than t^{-r} for any $r > 0$.

Consider now the advanced equation with positive and negative coefficients

$$\dot{x}(t) - \sum_{k=1}^{m}\left[a_k(t)x(h_k(t)) - b_k(t)x(g_k(t))\right] = 0, \quad t\geq 0. \tag{2.26}$$

to Contents, page 1

Theorem 2.13. *Suppose that $a_k(t)$ and $b_k(t)$ are Lebesgue measurable locally essentially bounded functions, $a_k(t) \geq b_k(t) \geq 0$, $h_k(t)$ and $g_k(t)$ are Lebesgue measurable functions, $h_k(t) \geq g_k(t) \geq t$, and inequality* (2.2) *has a nonnegative solution, then* (2.26) *has a nonoscillatory solution; moreover, it has a positive nondecreasing and a negative nonincreasing solutions.*

Proof. Let u_0 be a nonnegative solution of (2.2) and denote

$$u_{n+1}(t) = \sum_{k=1}^{m}\left(a_k(t)\exp\left\{\int_t^{h_k(t)} u_n(s)ds\right\} - b_k(t)\exp\left\{\int_t^{g_k(t)} u_n(s)ds\right\}\right), \quad t \geq t_0,\ n \geq 0. \tag{2.27}$$

We have $u_0 \geq 0$, and by (2.2),

$$\begin{aligned} u_0 &\geq \sum_{k=1}^{m} a_k(t)\exp\left\{\int_t^{h_k(t)} u_0(s)ds\right\} \\ &\geq \sum_{k=1}^{m}\left(a_k(t)\exp\left\{\int_t^{h_k(t)} u_0(s)ds\right\} - b_k(t)\exp\left\{\int_t^{g_k(t)} u_0(s)ds\right\}\right) = u_1(t). \end{aligned} \tag{2.28}$$

Since $a_k(t) \geq b_k(t) \geq 0$ and $t \leq g_k(t) \leq h_k(t)$, then $u_1(t) \geq 0$.

Next, let us assume that $0 \leq u_n(t) \leq u_{n-1}(t)$. The assumptions of the theorem imply $u_{n+1} \geq 0$. Let us demonstrate that $u_{n+1}(t) \leq u_n(t)$. This inequality has the form

$$\begin{aligned} &\sum_{k=1}^{m}\left(a_k(t)\exp\left\{\int_t^{h_k(t)} u_n(s)ds\right\} - b_k(t)\exp\left\{\int_t^{g_k(t)} u_n(s)ds\right\}\right) \\ &\quad \leq \sum_{k=1}^{m}\left(a_k(t)\exp\left\{\int_t^{h_k(t)} u_{n-1}(s)ds\right\} - b_k(t)\exp\left\{\int_t^{g_k(t)} u_{n-1}(s)ds\right\}\right), \end{aligned} \tag{2.29}$$

which is equivalent to

$$\begin{aligned} &\sum_{k=1}^{m}\exp\left\{\int_t^{h_k(t)} u_n(s)ds\right\}\left(a_k(t) - b_k(t)\exp\left\{-\int_{g_k(t)}^{h_k(t)} u_n(s)ds\right\}\right) \\ &\quad \leq \sum_{k=1}^{m}\exp\left\{\int_t^{h_k(t)} u_{n-1}(s)ds\right\}\left(a_k(t) - b_k(t)\exp\left\{-\int_{g_k(t)}^{h_k(t)} u_{n-1}(s)ds\right\}\right). \end{aligned} \tag{2.30}$$

This inequality is evident for any $0 \leq u_n(t) \leq u_{n-1}(t)$, $a_k(t) \geq 0$, and $b_k \geq 0$; thus, we have $u_{n+1}(t) \leq u_n(t)$.

By the Lebesgue convergence theorem, there is a pointwise limit $u(t) = \lim_{n\to\infty} u_n(t)$ satisfying

$$u(t) = \sum_{k=1}^{m}\left(a_k(t)\exp\left\{\int_t^{h_k(t)} u(s)ds\right\} - b_k(t)\exp\left\{\int_t^{g_k(t)} u(s)ds\right\}\right), \quad t \geq t_0, \tag{2.31}$$

$u(t) \geq 0$, $t \geq t_0$. Then, the function

$$x(t) = x(t_0) \exp\left\{ \int_{t_0}^{t} u(s) ds \right\}, \quad t \geq t_0 \tag{2.32}$$

is a positive nondecreasing solution of (2.26) for any $x(t_0) > 0$ and is a negative nonincreasing solution of (2.26) for any $x(t_0) < 0$. □

Corollary 2.14. *Let $a_k(t)$ and $b_k(t)$ be Lebesgue measurable locally essentially bounded functions satisfying $a_k(t) \geq b_k(t) \geq 0$, and let $h_k(t)$ and $g_k(t)$ be Lebesgue measurable functions, where $h_k(t) \geq g_k(t) \geq t$. Assume in addition that inequality* (2.7) *holds. Then,* (2.26) *has a nonoscillatory solution.*

Consider now the equation with constant deviations of advanced arguments

$$\dot{x}(t) - \sum_{k=1}^{m} [a_k(t)x(t+\tau_k) - b_k(t)x(t+\sigma_k)] = 0, \tag{2.33}$$

where a_k, b_k are continuous functions, $\tau_k \geq 0$, $\sigma_k \geq 0$.

Denote $A_k = \sup_{t\geq t_0} a_k(t)$, $a_k = \inf_{t\geq t_0} a_k(t)$, $B_k = \sup_{t\geq t_0} b_k(t)$, $b_k = \inf_{t\geq t_0} b_k(t)$.

Theorem 2.15. *Suppose that $a_k \geq 0$, $b_k \geq 0$, $A_k < \infty$, and $B_k < \infty$.*

If there exists a number $\lambda_0 < 0$ such that

$$\sum_{k=1}^{m} \left(a_k e^{\lambda_0 \tau_k} - B_k \right) \geq \lambda_0, \tag{2.34}$$

$$\sum_{k=1}^{m} \left(A_k - b_k e^{\lambda_0 \sigma_k} \right) \leq 0, \tag{2.35}$$

then (2.33) *has a nonoscillatory solution; moreover, it has a positive nonincreasing and a negative nondecreasing solutions.*

Proof. In the space $C[t_0,\infty)$, consider the set $M = \{u \mid \lambda_0 \leq u \leq 0\}$ and the operator

$$(Hu)(t) = \sum_{k=1}^{m} \left(a_k(t) \exp\left\{ \int_{t}^{t+\tau_k} u(s) ds \right\} - b_k(t) \exp\left\{ \int_{t}^{t+\sigma_k} u(s) ds \right\} \right). \tag{2.36}$$

For $u \in M$, we have from (2.34) and (2.35)

$$\begin{aligned} (Hu)(t) &\leq \sum_{k=1}^{m} \left(A_k - b_k e^{\lambda_0 \sigma_k} \right) \leq 0, \\ (Hu)(t) &\geq \sum_{k=1}^{m} \left(a_k e^{\lambda_0 \tau_k} - B_k \right) \geq \lambda_0. \end{aligned} \tag{2.37}$$

Hence, $HM \subset M$.

to Contents, page 1

Consider the integral operator

$$(Tu)(t) := \int_t^{t+\delta} u(s)ds, \quad \delta > 0. \tag{2.38}$$

We will show that TM is a compact set in the space $C[t_0, \infty)$. For $u \in M$, we have

$$\|(Tu)(t)\|_{C[t_0,\infty)} \le \sup_{t\ge t_0} \int_t^{t+\delta} |u(s)|ds \le |\lambda_0|\delta. \tag{2.39}$$

Hence, the functions in the set TM are uniformly bounded in the space $C[t_0, \infty)$.

The equality $\int_{t_0}^{t_0+\delta} - \int_t^{t+\delta} = \int_{t_0}^{t} + \int_t^{t_0+\delta} - \int_t^{t_0+\delta} - \int_{t_0+\delta}^{t+\delta} = \int_{t_0}^{t} - \int_{t_0+\delta}^{t+\delta}$ implies

$$\begin{aligned} |(Tu)(t) - (Tu)(t_0)| &= \left| \int_t^{t+\delta} u(s) - \int_{t_0}^{t_0+\delta} u(s)ds \right| \\ &\le \int_{t_0}^{t} |u(s)|ds + \int_{t_0+\delta}^{t+\delta} |u(s)|ds \le 2|\lambda_0||t - t_0|. \end{aligned} \tag{2.40}$$

Hence, the set TM and so the set HM are compact in the space $C[t_0, \infty)$.

By the Schauder fixed point theorem, there exists a continuous function u satisfying $\lambda_0 \le u \le 0$ such that $u = Hu$; thus, the function

$$x(t) = x(t_0) \exp\left\{ \int_{t_0}^{t} u(s)ds \right\}, \quad t \ge t_0 \tag{2.41}$$

is a positive nonincreasing solution of (2.33) for any $x(t_0) > 0$ and is a negative nondecreasing solution of (2.26) for any $x(t_0) < 0$. □

Let us remark that (2.35) for any $\lambda_0 < 0$ implies $\sum_{k=1}^{m}(A_k - b_k) < 0$.

Corollary 2.16. *Let* $\sum_{k=1}^{m}(A_k - b_k) < 0$, $\sum_{k=1}^{m} A_k > 0$, *and for*

$$\lambda_0 = \frac{\ln\left(\sum_{k=1}^{m} A_k / \sum_{k=1}^{m} b_k\right)}{\max_k \sigma_k}, \tag{2.42}$$

the inequality

$$\sum_{k=1}^{m} \left(a_k e^{\lambda_0 \tau_k} - B_k \right) \ge \lambda_0 \tag{2.43}$$

holds, then (2.33) *has a bounded positive solution.*

Proof. The negative number λ_0 defined in (2.42) is a solution of both (2.34) and (2.35); by definition, it satisfies (2.35), and (2.43) implies (2.34). □

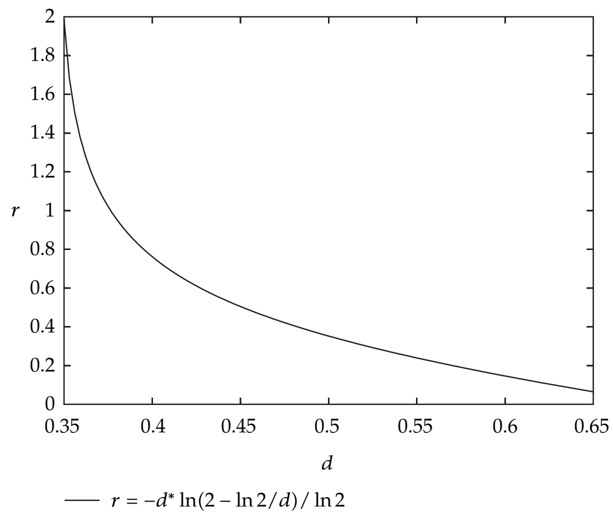


Figure 1: The domain of values (d, r) satisfying inequality (2.47). If the values of advances d and r are under the curve, then (2.44) has a positive solution.

Example 2.17. Consider the equation with constant advances and coefficients

$$\dot{x}(t) - ax(t+r) + bx(t+d) = 0, \tag{2.44}$$

where $0 < a < b$, $d > 0$, $r \geq 0$. Then, $\lambda_0 = (1/d)\ln(a/b)$ is the minimal value of λ for which inequality (2.35) holds; for (2.44), it has the form $a - be^{\lambda d} \leq 0$.

Inequality (2.34) for (2.44) can be rewritten as

$$f(\lambda) = ae^{\lambda r} - b - \lambda \geq 0, \tag{2.45}$$

where the function $f(x)$ decreases on $(-\infty, -\ln(ar)/r]$ if $\tau > 0$ and for any negative x if $r = 0$; besides, $f(0) < 0$. Thus, if $f(\lambda_1) < 0$ for some $\lambda_1 < 0$, then $f(\lambda) < 0$ for any $\lambda \in [\lambda_1, 0)$. Hence, the inequality

$$f(\lambda_0) = a\left(\frac{a}{b}\right)^{r/d} - b - \frac{1}{d}\ln\left(\frac{a}{b}\right) \geq 0 \tag{2.46}$$

is necessary and sufficient for the conditions of Theorem 2.15 to be satisfied for (2.44).

Figure 1 demonstrates possible values of advances d and r, such that Corollary 2.16 implies the existence of a positive solution in the case $1 = a < b = 2$. Then, (2.46) has the form $0.5^{r/d} \geq 2 - (\ln 2)/d$, which is possible only for $d > 0.5\ln 2 \approx 0.347$ and for these values is equivalent to

$$r \leq \frac{-d\ln(2 - \ln 2/d)}{\ln 2}. \tag{2.47}$$

to Contents, page 1

3. Comments and Open Problems

In this paper, we have developed nonoscillation theory for advanced equations with variable coefficients and advances. Most previous nonoscillation results deal with either oscillation or constant deviations of arguments. Among all cited papers, only [8] has a nonoscillation condition (Theorem 2.11) for a partial case of (2.1) (with $h_k(t) = t + \tau_k$), which in this case coincides with Corollary 2.4. The comparison of results of the present paper with the previous results of the authors was discussed in the introduction.

Finally, let us state some open problems and topics for research.

(1) Prove or disprove:
if (2.1), with $a_k(t) \geq 0$, has a nonoscillatory solution, then (2.26) with positive and negative coefficients also has a nonoscillatory solution.

As the first step in this direction, prove or disprove that if $h(t) \geq t$ and the equation

$$\dot{x}(t) - a^{+}(t)x(h(t)) = 0 \tag{3.1}$$

has a nonoscillatory solution, then the equation

$$\dot{x}(t) - a(t)x(h(t)) = 0 \tag{3.2}$$

also has a nonoscillatory solution, where $a^{+}(t) = \max\{a(t), 0\}$.

If these conjectures are valid, obtain comparison results for advanced equations.

(2) Deduce nonoscillation conditions for (2.1) with oscillatory coefficients. Oscillation results for an equation with a constant advance and an oscillatory coefficient were recently obtained in [22].

(3) Consider advanced equations with positive and negative coefficients when the numbers of positive and negative terms do not coincide.

(4) Study existence and/or uniqueness problem for the initial value problem or boundary value problems for advanced differential equations.

Acknowledgments

L. Berezansky was partially supported by the Israeli Ministry of Absorption. E. Braverman was partially supported by NSERC.

References

[1] D. M. Dubois, "Extension of the Kaldor-Kalecki model of business cycle with a computational anticipated capital stock," *Journal of Organisational Transformation and Social Change*, vol. 1, no. 1, pp. 63–80, 2004.
[2] A. Kaddar and H. Talibi Alaoui, "Fluctuations in a mixed IS-LM business cycle model," *Electronic Journal of Differential Equations*, vol. 2008, no. 134, pp. 1–9, 2008.
[3] G. Ladas and I. P. Stavroulakis, "Oscillations caused by several retarded and advanced arguments," *Journal of Differential Equations*, vol. 44, no. 1, pp. 134–152, 1982.

[4] Y. Kitamura and T. Kusano, "Oscillation of first-order nonlinear differential equations with deviating arguments," *Proceedings of the American Mathematical Society*, vol. 78, no. 1, pp. 64–68, 1980.

[5] L. H. Erbe, Q. Kong, and B. G. Zhang, *Oscillation Theory for Functional-Differential Equations*, vol. 190 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 1995.

[6] N. T. Markova and P. S. Simeonov, "Oscillation criteria for first order nonlinear differential equations with advanced arguments," *Communications in Applied Analysis*, vol. 10, no. 2-3, pp. 209–221, 2006.

[7] R. P. Agarwal, S. R. Grace, and D. O'Regan, "On the oscillation of certain advanced functional differential equations using comparison methods," *Fasciculi Mathematici*, no. 35, pp. 5–22, 2005.

[8] X. Li and D. Zhu, "Oscillation and nonoscillation of advanced differential equations with variable coefficients," *Journal of Mathematical Analysis and Applications*, vol. 269, no. 2, pp. 462–488, 2002.

[9] Q. Meng and J. Yan, "Nonautonomous differential equations of alternately retarded and advanced type," *International Journal of Mathematics and Mathematical Sciences*, vol. 26, no. 10, pp. 597–603, 2001.

[10] L. Dai and C. Tian, "Oscillation criteria for advanced type differential equations," *Journal of Ocean University of Qingdao*, vol. 30, no. 2, pp. 364–368, 2000.

[11] I.-G. E. Kordonis and Ch. G. Philos, "Oscillation and nonoscillation in delay or advanced differential equations and in integrodifferential equations," *Georgian Mathematical Journal*, vol. 6, no. 3, pp. 263–284, 1999.

[12] B. Baculíková, R. P. Agarwal, T. Li, and J. Džurina, "Oscillation of third-order nonlinear functional differential equations with mixed arguments," to appear in *Acta Mathematica Hungarica*.

[13] I. Győri and G. Ladas, *Oscillation Theory of Delay Differential Equations*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, NY, USA, 1991.

[14] T. Kusano, "On even-order functional-differential equations with advanced and retarded arguments," *Journal of Differential Equations*, vol. 45, no. 1, pp. 75–84, 1982.

[15] R. Koplatadze, "On asymptotic behavior of solutions of n-th order Emden-Fowler differential equations with advanced argument," *Czechoslovak Mathematical Journal*, vol. 60(135), no. 3, pp. 817–833, 2010.

[16] S. Murugadass, E. Thandapani, and S. Pinelas, "Oscillation criteria for forced second-order mixed type quasilinear delay differential equations," *Electronic Journal of Differential Equations*, vol. 2010, no. 73, pp. 1–9, 2010.

[17] A. Zafer, "Interval oscillation criteria for second order super-half linear functional differential equations with delay and advanced arguments," *Mathematische Nachrichten*, vol. 282, no. 9, pp. 1334–1341, 2009.

[18] S. R. Grace, R. P. Agarwal, R. Pavani, and E. Thandapani, "On the oscillation of certain third order nonlinear functional differential equations," *Applied Mathematics and Computation*, vol. 202, no. 1, pp. 102–112, 2008.

[19] D. R. Anderson, "Oscillation of second-order forced functional dynamic equations with oscillatory potentials," *Journal of Difference Equations and Applications*, vol. 13, no. 5, pp. 407–421, 2007.

[20] L. Berezansky and Y. Domshlak, "Differential equations with several deviating arguments: Sturmian comparison method in oscillation theory. II," *Electronic Journal of Differential Equations*, vol. 2002, no. 31, pp. 1–18, 2002.

[21] L. Berezansky, E. Braverman, and S. Pinelas, "On nonoscillation of mixed advanced-delay differential equations with positive and negative coefficients," *Computers & Mathematics with Applications*, vol. 58, no. 4, pp. 766–775, 2009.

[22] X. Li, D. Zhu, and H. Wang, "Oscillation for advanced differential equations with oscillating coefficients," *International Journal of Mathematics and Mathematical Sciences*, no. 33, pp. 2109–2118, 2003.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 326052, 22 pages
doi:10.1155/2011/326052

Research Article

On Nonseparated Three-Point Boundary Value Problems for Linear Functional Differential Equations

A. Rontó[1] and M. Rontó[2]

[1] *Institute of Mathematics, Academy of Sciences of the Czech Republic, 22 Žižkova St., 61662 Brno, Czech Republic*
[2] *Department of Analysis, University of Miskolc, 3515 Miskolc-Egyetemváros, Hungary*

Correspondence should be addressed to A. Rontó, ronto@math.cas.cz

Received 20 January 2011; Accepted 27 April 2011

Academic Editor: Yuri V. Rogovchenko

For a system of linear functional differential equations, we consider a three-point problem with nonseparated boundary conditions determined by singular matrices. We show that, to investigate such a problem, it is often useful to reduce it to a parametric family of two-point boundary value problems for a suitably perturbed differential system. The auxiliary parametrised two-point problems are then studied by a method based upon a special kind of successive approximations constructed explicitly, whereas the values of the parameters that correspond to solutions of the original problem are found from certain numerical determining equations. We prove the uniform convergence of the approximations and establish some properties of the limit and determining functions.

1. Introduction

The aim of this paper is to show how a suitable parametrisation can help when dealing with nonseparated three-point boundary conditions determined by singular matrices. We construct a suitable numerical-analytic scheme allowing one to approach a three-point boundary value problem through a certain iteration procedure. To explain the term, we recall that, formally, the methods used in the theory of boundary value problems can be characterised as analytic, functional-analytic, numerical, or numerical-analytic ones.

While the analytic methods are generally used for the investigation of qualitative properties of solutions such as the existence, multiplicity, branching, stability, or dichotomy and generally use techniques of calculus (see, e.g., [1–11] and the references in [12]), the functional-analytic ones are based mainly on results of functional analysis and topological

to Contents, page 1

degree theory and essentially use various techniques related to operator equations in abstract spaces [13–26]. The numerical methods, under the assumption on the existence of solutions, provide practical numerical algorithms for their approximation [27, 28]. The numerical construction of approximate solutions is usually based on an idea of the shooting method and may face certain difficulties because, as a rule, this technique requires some global regularity conditions, which, however, are quite often satisfied only locally.

Methods of the so-called numerical-analytic type, in a sense, combine, advantages of the mentioned approaches and are usually based upon certain iteration processes constructed explicitly. Such an approach belongs to the few of them that offer constructive possibilities both for the investigation of the existence of a solution and its approximate construction. In the theory of nonlinear oscillations, numerical-analytic methods of this kind had apparently been first developed in [20, 29–31] for the investigation of periodic boundary value problems. Appropriate versions were later developed for handling more general types of nonlinear boundary value problems for ordinary and functional-differential equations. We refer, for example, to the books [12, 32–34], the handbook [35], the papers [36–50], and the survey [51–57] for related references.

For a boundary value problem, the numerical-analytic approach usually replaces the problem by the Cauchy problem for a suitably perturbed system containing some artificially introduced vector parameter z, which most often has the meaning of an initial value of the solution and the numerical value of which is to be determined later. The solution of Cauchy problem for the perturbed system is sought for in an analytic form by successive approximations. The functional "perturbation term," by which the modified equation differs from the original one, depends explicitly on the parameter z and generates a system of algebraic or transcendental "determining equations" from which the numerical values of z should be found. The solvability of the determining system, in turn, may by checked by studying some of its approximations that are constructed explicitly.

For example, in the case of the two-point boundary value problem

$$x'(t) = f(t, x(t)), \quad t \in [a, b], \tag{1.1}$$

$$Ax(a) + Dx(b) = d, \tag{1.2}$$

where $x : [a, b] \to \mathbb{R}^n, -\infty < a < b < +\infty$, $d \in \mathbb{R}^n$, $\det D \neq 0$, the corresponding Cauchy problem for the modified parametrised system of integrodifferential equations has the form [12]

$$\begin{gathered} x'(t) = f(t, x(t)) + \frac{1}{b-a}\left(D^{-1}d - \left(D^{-1}A + \mathbb{1}_n\right)z\right) - \frac{1}{b-a}\int_a^b f(s, x(s))ds, \quad t \in [a, b], \\ x(a) = z, \end{gathered} \tag{1.3}$$

where $\mathbb{1}_n$ is the unit matrix of dimension n and the parameter $z \in \mathbb{R}^n$ has the meaning of initial value of the solution at the point a. The expression

$$\frac{1}{b-a}\left(D^{-1}d - \left(D^{-1}A + \mathbb{1}_n\right)z\right) - \frac{1}{b-a}\int_a^b f(s, x(s))ds \tag{1.4}$$

to Contents, page 1

in (1.3) plays the role of a "perturbation term" and its choice is, of course, not unique. The solution of problem (1.3) is sought for in an analytic form by the method of successive approximations similar to the Picard iterations. According to the formulas

$$x_{m+1}(t,z) := z + \int_a^t \left(f(s, x_m(s,z))ds - \frac{1}{b-a}\int_a^b f(\tau, x_m(\tau,z))d\tau \right) ds \\ + \frac{t-a}{b-a}\left(D^{-1}d - \left(D^{-1}A + \mathbb{1}_n\right)z\right), \quad m = 0,1,2,\ldots, \tag{1.5}$$

where $x_0(t,z) := z$ for all $t \in [a,b]$ and $z \in \mathbb{R}^n$, one constructs the iterations $x_m(\cdot,z)$, $m = 1,2,\ldots$, which depend upon the parameter z and, for arbitrary its values, satisfy the given boundary conditions (1.2): $Ax_m(a,z) + Dx_m(b,z) = d$, $z \in \mathbb{R}^n$, $m = 1,2,\ldots$. Under suitable assumptions, one proves that sequence (1.5) converges to a limit function $x_\infty(\cdot,z)$ for any value of z.

The procedure of passing from the original differential system (1.1) to its "perturbed" counterpart and the investigation of the latter by using successive approximations (1.5) leads one to the system of determining equations

$$D^{-1}d - \left(D^{-1}A + \mathbb{1}_n\right)z - \int_a^b f(s, x_\infty(s,z))ds = 0, \tag{1.6}$$

which gives those numerical values $z = z_*$ of the parameter that correspond to solutions of the given boundary value problem (1.1), (1.10). The form of system (1.6) is, of course, determined by the choice of the perturbation term (1.4); in some other related works, auxiliary equations are constructed in a different way (see, e.g., [42]). It is clear that the complexity of the given equations and boundary conditions has an essential influence both on the possibility of an efficient construction of approximate solutions and the subsequent solvability analysis.

The aim of this paper is to extend the techniques used in [46] for the system of n linear functional differential equations of the form

$$x'(t) = P_0(t)x(t) + P_1(t)x(\beta(t)) + f(t), \quad t \in [0,T], \tag{1.7}$$

subjected to the inhomogeneous three-point Cauchy-Nicoletti boundary conditions

$$\begin{gathered} x_1(0) = x_{10},\ldots,\ x_p(0) = x_{p0}, \\ x_{p+1}(\xi) = d_{p+1},\ldots,\ x_{p+q}(\xi) = d_{p+q}, \\ x_{p+q+1}(T) = d_{p+q+1},\ldots,\ x_n(T) = d_n, \end{gathered} \tag{1.8}$$

with $\xi \in (0,T)$ is given and $x = \mathrm{col}(x_1,\ldots,x_n)$, to the case where the system of linear functional differential equations under consideration has the general form

$$x'(t) = (lx)(t) + f(t), \quad t \in [a,b], \tag{1.9}$$

to Contents, page 1

and the three-point boundary conditions are non-separated and have the form

$$Ax(a) + Bx(\xi) + Cx(b) = d, \tag{1.10}$$

where A, B, and C are singular matrices, $d = \operatorname{col}(d_1, \ldots, d_n)$. Here, $l = (l_k)_{k=1}^n : C([a,b], \mathbb{R}^n) \to L_1([a,b], \mathbb{R}^n)$ is a bounded linear operator and $f \in L_1([a,b], \mathbb{R}^n)$ is a given function.

It should be noted that, due to the singularity of the matrices that determine the boundary conditions (1.10), certain technical difficulties arise which complicate the construction of successive approximations.

The following notation is used in the sequel:

$C([a,b], \mathbb{R}^n)$ is the Banach space of the continuous functions $[a,b] \to \mathbb{R}^n$ with the standard uniform norm;

$L_1([a,b], \mathbb{R}^n)$ is the usual Banach space of the vector functions $[a,b] \to \mathbb{R}^n$ with Lebesgue integrable components;

$\mathcal{L}(\mathbb{R}^n)$ is the algebra of all the square matrices of dimension n with real elements;

$r(Q)$ is the maximal, in modulus, eigenvalue of a matrix $Q \in \mathcal{L}(\mathbb{R}^n)$;

$\mathbb{1}_k$ is the unit matrix of dimension k;

$\mathbb{0}_{i,j}$ is the zero matrix of dimension $i \times j$;

$\mathbb{0}_i = \mathbb{0}_{i,i}$.

2. Problem Setting and Freezing Technique

We consider the system of n linear functional differential equations (1.9) subjected to the nonseparated inhomogeneous three-point boundary conditions of form (1.10). In the boundary value problem (1.1), (1.10), we suppose that $-\infty < a < b < \infty$, $l = (l_k)_{k=1}^n : C([a,b], \mathbb{R}^n) \to L_1([a,b], \mathbb{R}^n)$ is a bounded linear operator, $f : [a,b] \to \mathbb{R}^n$ is an integrable function, $d \in \mathbb{R}^n$ is a given vector, A, B, and C are singular square matrices of dimension n, and C has the form

$$C = \begin{pmatrix} V & W \\ \mathbb{0}_{n-q,q} & \mathbb{0}_{n-q} \end{pmatrix}, \tag{2.1}$$

where V is nonsingular square matrix of dimension $q < n$ and W is an arbitrary matrix of dimension $q \times (n-q)$. The singularity of the matrices determining the boundary conditions (1.10) causes certain technical difficulties. To avoid dealing with singular matrices in the boundary conditions and simplify the construction of a solution in an analytic form, we use a two-stage parametrisation technique. Namely, we first replace the three-point boundary conditions by a suitable parametrised family of two-point inhomogeneous conditions, after which one more parametrisation is applied in order to construct an auxiliary perturbed differential system. The presence of unknown parameters leads one to a certain system of determining equations, from which one finds those numerical values of the parameters that correspond to the solutions of the given three-point boundary value problem.

to Contents, page 1

We construct the auxiliary family of two-point problems by "freezing" the values of certain components of x at the points ξ and b as follows:

$$\operatorname{col}(x_1(\xi),\ldots,x_n(\xi))=\lambda,$$
$$\operatorname{col}\big(x_{q+1}(b),\ldots,x_n(b)\big)=\eta, \tag{2.2}$$

where $\lambda=\operatorname{col}(\lambda_1,\ldots,\lambda_n)\in\mathbb{R}^n$ and $\eta=col(\eta_1,\ldots,\eta_{n-q})\in\mathbb{R}^{n-q}$ are vector parameters. This leads us to the parametrised two-point boundary condition

$$Ax(a)+Dx(b)=d-B\lambda-N_q\eta, \tag{2.3}$$

where

$$N_q:=\begin{pmatrix}\mathbb{0}_{q,n-q}\\ \mathbb{1}_{n-q}\end{pmatrix} \tag{2.4}$$

and the matrix D is given by the formula

$$D:=\begin{pmatrix}V & W\\ \mathbb{0}_{n-q,q} & \mathbb{1}_{n-q}\end{pmatrix} \tag{2.5}$$

with a certain rectangular matrix W of dimension $q\times(n-q)$. It is important to point out that the matrix D appearing in the two-point condition (2.3) is non-singular.

It is easy to see that the solutions of the original three-point boundary value problem (1.1), (1.10) coincide with those solutions of the two-point boundary value problem (1.1), (2.3) for which the additional condition (2.2) is satisfied.

Remark 2.1. The matrices A and B in the boundary conditions (1.10) are arbitrary and, in particular, may be singular. If the number r of the linearly independent boundary conditions in (1.10) is less than n, that is, the rank of the $(n\times 3n)$-dimensional matrix $[A,B,C]$ is equal to r, then the boundary value problem (1.1), (1.10) may have an $(n-r)$-parametric family of solutions.

We assume that throughout the paper the operator l determining the system of equations (1.9) is represented in the form

$$l=l^0-l^1, \tag{2.6}$$

where $l^j=(l_k^j)_{k=1}^n : C([a,b],\mathbb{R}^n)\to L_1([a,b],\mathbb{R}^n)$, $j=0,1$, are bounded linear operators posi-tive in the sense that $(l_k^j u)(t)\geq 0$ for a.e. $t\in[a,b]$ and any $k=1,2,\ldots,n$, $j=0,1$, and $u\in C([a,b],\mathbb{R}^n)$ such that $\min_{t\in[a,b]}u_k(t)\geq 0$ for all $k=1,2,\ldots,n$. We also put $\widehat{l}_k:=l_k^0+l_k^1$, $k=1,2,\ldots,n$, and

$$\widehat{l}:=l^0+l^1. \tag{2.7}$$

to Contents, page 1

3. Auxiliary Estimates

In the sequel, we will need several auxiliary statements.

Lemma 3.1. *For an arbitrary essentially bounded function $u : [a,b] \to \mathbb{R}$, the estimates*

$$\left| \int_a^t \left(u(\tau) - \frac{1}{b-a} \int_a^b u(s)ds \right) d\tau \right| \leq \alpha(t) \left(\operatorname*{ess\,sup}_{s\in[a,b]} u(s) - \operatorname*{ess\,inf}_{s\in[a,b]} u(s) \right) \tag{3.1}$$

$$\leq \frac{b-a}{4} \left(\operatorname*{ess\,sup}_{s\in[a,b]} u(s) - \operatorname*{ess\,inf}_{s\in[a,b]} u(s) \right) \tag{3.2}$$

are true for all $t \in [a,b]$, where

$$\alpha(t) := (t-a)\left(1 - \frac{t-a}{b-a}\right), \quad t \in [a,b]. \tag{3.3}$$

Proof. Inequality (3.1) is established similarly to [58, Lemma 3] (see also [12, Lemma 2.3]), whereas (3.2) follows directly from (3.1) if the relation

$$\max_{t\in[a,b]} \alpha(t) = \frac{1}{4}(b-a) \tag{3.4}$$

is taken into account. □

Let us introduce some notation. For any $k = 1,2,\ldots,n$, we define the n-dimensional row-vector e_k by putting

$$e_k := (\underbrace{0,0,\ldots,0,}_{k-1}1,0,\ldots,0). \tag{3.5}$$

Using operators (2.7) and the unit vectors (3.5), we define the matrix-valued function $K_l : [a,b] \to \mathcal{L}(\mathbb{R}^n)$ by setting

$$K_l := \left[\widehat{l}e_1^*, \widehat{l}e_2^*, \ldots, \widehat{l}e_n^*\right]. \tag{3.6}$$

Note that, in (3.6), $\widehat{l}e_i^*$ means the value of the operator $\widehat{l}$ on the constant vector function is equal identically to e_i^*, where e_i^* is the vector transpose to e_i. It is easy to see that the components of K_l are Lebesgue integrable functions.

Lemma 3.2. *The componentwise estimate*

$$|(lx)(t)| \leq K_l(t) \max_{s\in[a,b]} |x(s)|, \quad t \in [a,b], \tag{3.7}$$

is true for any $x \in C([a,b],\mathbb{R}^n)$, where $K_l : [a,b] \to \mathcal{L}(\mathbb{R}^n)$ is the matrix-valued function given by formula (3.6).

to Contents, page 1

Proof. Let $x = (x_k)_{k=1}^{n}$ be an arbitrary function from $C([a,b],\mathbb{R}^n)$. Then, recalling the notation for the components of l, we see that

$$lx = \sum_{i=1}^{n} e_i^* l_i x. \tag{3.8}$$

On the other hand, due to (3.5), we have $x = \sum_{k=1}^{n} e_k^* x_k$ and, therefore, by virtue of (3.8) and (2.6),

$$lx = \sum_{i=1}^{n} e_i^* l_i x = \sum_{i=1}^{n} e_i^* l_i \left(\sum_{k=1}^{n} e_k^* x_k \right) = \sum_{i=1}^{n} e_i^* \left(\sum_{k=1}^{n} \left(l_i^0 e_k^* x_k - l_i^1 e_k^* x_k \right) \right). \tag{3.9}$$

On the other hand, the obvious estimate

$$\sigma x_k(t) \le \max_{s\in[a,b]} |x_k(s)|, \quad t \in [a,b],\ k = 1,2,\ldots,n,\ \sigma \in \{-1,1\}, \tag{3.10}$$

and the positivity of the operators l^j, $j = 0,1$, imply

$$l_i^j(\sigma x_k)(t) = \sigma\left(l_i^j x_k\right)(t) \le l_i^j \max_{s\in[a,b]} |x_k(s)| \tag{3.11}$$

for a.e. $t \in [a,b]$ and any $k, j = 1,2,\ldots,n$, $\sigma \in \{-1,1\}$. This, in view of (2.7) and (3.9), leads us immediately to estimate (3.7). □

4. Successive Approximations

To study the solution of the auxiliary two-point parametrised boundary value problem (1.9), (2.3) let us introduce the sequence of functions $x_m : [a,b] \times \mathbb{R}^{3n-q} \to \mathbb{R}^n$, $m \ge 0$, by putting

$$\begin{aligned} x_{m+1}(t,z,\lambda,\eta) &:= \varphi_{z,\lambda,\eta}(t) + \int_a^t ((lx_m(\cdot,z,\lambda,\eta))(s) + f(s))ds \\ &\quad - \frac{t-a}{b-a}\int_a^b ((lx_m(\cdot,z,\lambda,\eta))(s) + f(s))ds, \quad m = 0,1,2,\ldots, \\ x_0(t,z,\lambda,\eta) &:= \varphi_{z,\lambda,\eta}(t) \end{aligned} \tag{4.1}$$

for all $t \in [a,b]$, $z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, and $\eta \in \mathbb{R}^{n-q}$, where

$$\varphi_{z,\lambda,\eta}(t) := z + \frac{t-a}{b-a}\left(D^{-1}(d - B\lambda + N_q\eta) - \left(D^{-1}A + \mathbb{1}_n\right)z\right). \tag{4.2}$$

In the sequel, we consider x_m as a function of t and treat the vectors z, λ, and η as parameters.

to Contents, page 1

Lemma 4.1. *For any $m \geq 0$, $t \in [a,b]$, $z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, and $\eta \in \mathbb{R}^{n-q}$, the equalities*

$$\begin{aligned} x_m(a,z,\lambda,\eta) &= z, \\ Ax_m(a,z,\lambda,\eta) + Dx_m(b,z,\lambda,\eta) &= d - B\lambda + N_q\eta, \end{aligned} \tag{4.3}$$

are true.

The proof of Lemma 4.1 is carried out by straightforward computation. We emphasize that the matrix D appearing in the two-point condition (2.3) is non-singular. Let us also put

$$(\mathcal{M}y)(t) := \left(1 - \frac{t-a}{b-a}\right)\int_a^t y(s)ds + \frac{t-a}{b-a}\int_t^b y(s)ds, \quad t \in [a,b], \tag{4.4}$$

for an arbitrary $y \in L_1([a,b],\mathbb{R}^n)$. It is clear that $\mathcal{M} : L_1([a,b],\mathbb{R}^n) \to C([a,b],\mathbb{R}^n)$ is a positive linear operator. Using the operator $\mathcal{M}$, we put

$$Q_l := \left[\mathcal{M}(K_l e_1^*), \mathcal{M}(K_l e_2^*), \ \ldots, \mathcal{M}(K_l e_n^*)\right], \tag{4.5}$$

where K_l is given by formula (3.6). Finally, define a constant square matrix Q_l of dimension n by the formula

$$Q_l := \max_{t \in [a,b]} Q_l(t). \tag{4.6}$$

We point out that, as before, the maximum in (4.6) is taken componentwise (one should remember that, for $n > 1$, a point $t_* \in [a,b]$ such that $Q_l = Q_l(t_*)$ may not exist).

Theorem 4.2. *If the spectral radius of the matrix Q_l satisfies the inequality*

$$r(Q_l) < 1, \tag{4.7}$$

then, for arbitrary fixed $z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, and $\eta \in \mathbb{R}^{n-q}$:

(1) *the sequence of functions* (4.1) *converges uniformly in $t \in [a,b]$ for any fixed $(z,\lambda,\eta) \in \mathbb{R}^{3n-q}$ to a limit function*

$$x_\infty(t,z,\lambda,\eta) = \lim_{m\to+\infty} x_m(t,z,\lambda,\eta); \tag{4.8}$$

(2) *the limit function $x_\infty(\cdot,z,\lambda,\eta)$ possesses the properties*

$$\begin{aligned} x_\infty(a,z,\lambda,\eta) &= z, \\ Ax_\infty(a,z,\lambda,\eta) + Dx_\infty(b,z,\lambda,\eta) &= d - B\lambda + N_q\eta; \end{aligned} \tag{4.9}$$

to Contents, page 1

(3) *function* (4.8) *is a unique absolutely continuous solution of the integro-functional equation*

$$x(t) = z + \frac{t-a}{b-a}\left(D^{-1}(d - B\lambda + N_q\eta) - \left(D^{-1}A + \mathbb{1}_n\right)z\right) + \int_a^t ((lx)(s) + f(s))ds - \frac{t-a}{b-a}\int_a^b ((lx)(s) + f(s))ds, \quad t \in [a,b]; \tag{4.10}$$

(4) *the error estimate*

$$\max_{t\in[a,b]}\left|x_\infty(t,z,\lambda,\eta) - x_m(t,z,\lambda,\eta)\right| \le \frac{b-a}{4}Q_l^m(\mathbb{1}_n - Q_l)^{-1}\omega(z,\lambda,\eta) \tag{4.11}$$

holds, where $\omega : \mathbb{R}^{3n-q} \to \mathbb{R}^n$ *is given by the equality*

$$\omega(z,\lambda,\eta) := \operatorname*{ess\,sup}_{s\in[a,b]}\left((l\varphi_{z,\lambda,\eta})(s) + f(s)\right) - \operatorname*{ess\,inf}_{s\in[a,b]}\left((l\varphi_{z,\lambda,\eta})(s) + f(s)\right). \tag{4.12}$$

In (3.6), (4.11) and similar relations, the signs $|\cdot|$, $\le$, $\ge$, as well as the operators "max", "ess sup", "ess inf", and so forth, applied to vectors or matrices are understood componentwise.

Proof. The validity of assertion 1 is an immediate consequence of the formula (4.1). To obtain the other required properties, we will show, that under the conditions assumed, sequence (4.1) is a Cauchy sequence in the Banach space $C([a,b],\mathbb{R}^n)$ equipped with the standard uniform norm. Let us put

$$r_m(t,z,\lambda,\eta) := x_{m+1}(t,z,\lambda,\eta) - x_m(t,z,\lambda,\eta) \tag{4.13}$$

for all $z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, $\eta \in \mathbb{R}^{n-q}$, $t \in [a,b]$, and $m \ge 0$. Using Lemma 3.2 and taking equality (3.4) into account, we find that (4.1) yields

$$\begin{aligned}\left|x_1(t,z,\lambda,\eta) - x_0(t,z,\lambda,\eta)\right| &= \left|\int_a^t \left[(l\varphi_{z,\lambda,\eta})(s) + f(s)\right]ds - \frac{t-a}{b-a}\int_a^b \left[(l\varphi_{z,\lambda,\eta})(s) + f(s)\right]ds\right| \\ &\le \alpha(t)\omega(z,\lambda,\eta) \\ &\le \frac{b-a}{4}\omega(z,\lambda,\eta),\end{aligned} \tag{4.14}$$

for arbitrary fixed z, λ, and η, where α is the function given by (3.3) and $\omega(\cdot)$ is defined by formula (4.12).

to Contents, page 1

According to formulae (4.1), for all $t \in [a,b]$, arbitrary fixed z, λ, and η and $m = 1,2,\ldots$ we have

$$\begin{aligned} r_m(t,z,\lambda,\eta) &= \int_a^t l(x_m(\cdot,z,\lambda,\eta) - x_{m-1}(\cdot,z,\lambda,\eta))(s)ds \\ &\quad - \frac{t-a}{b-a}\int_a^b l(x_m(\cdot,z,\lambda,\eta) - x_{m-1}(\cdot,z,\lambda,\eta))(s)ds \\ &= \left(1 - \frac{t-a}{b-a}\right)\int_a^t l(x_m(\cdot,z,\lambda,\eta) - x_{m-1}(\cdot,z,\lambda,\eta))(s)ds \\ &\quad - \frac{t-a}{b-a}\int_t^b l(x_m(\cdot,z,\lambda,\eta) - x_{m-1}(\cdot,z,\lambda,\eta))(s)ds. \end{aligned} \tag{4.15}$$

Equalities (4.13) and (4.15) imply that for all $m = 1,2,\ldots$, arbitrary fixed z,λ,η and $t \in [a,b]$,

$$\begin{aligned} |r_m(t,z,\lambda,\eta)| &\le \left(1 - \frac{t-a}{b-a}\right)\int_a^t |l\ (r_{m-1}(\cdot,z,\lambda,\eta)(s))|ds \\ &\quad + \frac{t-a}{b-a}\int_t^b |l(r_{m-1}(\cdot,z,\lambda,\eta))(s)|ds. \end{aligned} \tag{4.16}$$

Applying inequality (3.7) of Lemma 3.2 and recalling formulae (4.5) and (4.6), we get

$$\begin{aligned} |r_m(t,z,\lambda,\eta)| &\le \left(1 - \frac{t-a}{b-a}\right)\int_a^t K_l(s)\max_{\tau\in[a,b]}|r_{m-1}(\tau,z,\lambda,\eta)|ds \\ &\quad + \frac{t-a}{b-a}\int_t^b K_l(s)\max_{\tau\in[a,b]}|r_{m-1}(\tau,z,\lambda,\eta)|ds \\ &= \left(\left(1 - \frac{t-a}{b-a}\right)\int_a^t K_l(s)ds + \frac{t-a}{b-a}\int_t^b K_l(s)ds\right)\max_{\tau\in[a,b]}|r_{m-1}(\tau,z,\lambda,\eta)| \\ &= Q_l(t)\max_{\tau\in[a,b]}|r_{m-1}(\tau,z,\lambda,\eta)| \\ &\le Q_l\max_{\tau\in[a,b]}|r_{m-1}(\tau,z,\lambda,\eta)|. \end{aligned} \tag{4.17}$$

Using (4.17) recursively and taking (4.14) into account, we obtain

$$\begin{aligned} |r_m(t,z,\lambda,\eta)| &\le Q_l^m \max_{\tau\in[a,b]}|r_0(\tau,z,\lambda,\eta)| \\ &\le \frac{b-a}{4}Q_l^m\omega(z,\lambda,\eta), \end{aligned} \tag{4.18}$$

to Contents, page 1

for all $m \geq 1$, $t \in [a,b]$, $z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, and $\eta \in \mathbb{R}^{n-q}$. Using (4.18) and (4.13), we easily obtain that, for an arbitrary $j \in \mathbb{N}$,

$$\begin{aligned}
\left|x_{m+j}(t,z,\lambda,\eta) - x_m(t,z,\lambda,\eta)\right| &= \big|\left(x_{m+j}(t,z,\lambda,\eta) - x_{m+j-1}(t,z,\lambda,\eta)\right) \\
&\quad + \left(x_{m+j-1}(t,z,\lambda,\eta) - x_{m+j-2}(t,z,\lambda,\eta)\right) + \cdots \\
&\quad + \left(x_{m+1}(t,z,\lambda,\eta) - x_{\mathrm{m}}(t,z,\lambda,\eta)\right)\big| \\
&\leq \sum_{i=0}^{j-1} \left|r_{m+i}(t,z,\lambda,\eta)\right| \\
&\leq \frac{b-a}{4} \sum_{i=0}^{j-1} Q_l^{m+i} \omega(z,\lambda,\eta).
\end{aligned} \tag{4.19}$$

Therefore, by virtue of assumption (4.7), it follows that

$$\begin{aligned}
\left|x_{m+j}(t,z,\lambda,\eta) - x_m(t,z,\lambda,\eta)\right| &\leq \frac{b-a}{4} Q_l^m \sum_{i=0}^{+\infty} Q_l^i \omega(z,\lambda,\eta) \\
&= \frac{b-a}{4} Q_l^m (\mathbb{1}_n - Q_l)^{-1} \omega(z,\lambda,\eta)
\end{aligned} \tag{4.20}$$

for all $m \geq 1$, $j \geq 1$, $t \in [a,b]$, $z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, and $\eta \in \mathbb{R}^{n-q}$. We see from (4.20) that (4.1) is a Cauchy sequence in the Banach space $C([a,b],\mathbb{R}^n)$ and, therefore, converges uniformly in $t \in [a,b]$ for all $(z,\lambda,\eta) \in \mathbb{R}^{3n-q}$:

$$\lim_{m\to\infty} x_m(t,z,\lambda,\eta) = x_\infty(t,z,\lambda,\eta), \tag{4.21}$$

that is, assertion 2 holds. Since all functions $x_m(t,z,\lambda,\eta)$ of the sequence (4.1) satisfy the boundary conditions (2.3), by passing to the limit in (2.3) as $m \to +\infty$ we show that the function $x_\infty(\cdot,z,\lambda,\eta)$ satisfies these conditions.

Passing to the limit as $m \to \infty$ in (4.1), we show that the limit function is a solution of the integro-functional equation (4.10). Passing to the limit as $j \to \infty$ in (4.20) we obtain the estimate

$$\left|x_\infty(t,z,\lambda,\eta) - x_m(t,z,\lambda,\eta)\right| \leq \frac{b-a}{4} Q_l^m (\mathbb{1}_n - Q_l)^{-1} \omega(z,\lambda,\eta) \tag{4.22}$$

for a.e. $t \in [a,b]$ and arbitrary fixed z, λ, η, and $m = 1,2,\ldots$. This completes the proof of Theorem 4.2. □

We have the following simple statement.

to Contents, page 1

Proposition 4.3. *If, under the assumptions of Theorem 4.2, one can specify some values of z, λ, and η, such that the limit function $x_\infty(\cdot, z, \lambda, \eta)$ possesses the property*

$$D^{-1}(d - B\lambda + N_q\eta) - \left(D^{-1}A + \mathbb{1}_n\right)z = \int_a^b ((lx_\infty(\cdot, z, \lambda, \eta))(s) + f(s))ds = 0, \tag{4.23}$$

then, for these $z, \lambda,$ and η, it is also a solution of the boundary value problem (1.9), (2.3).

Proof. The proof is a straightforward application of the above theorem. □

5. Some Properties of the Limit Function

Let us first establish the relation of the limit function $x_\infty(\cdot, z, \lambda, \eta)$ to the auxiliary two-point boundary value problem (1.9), (2.3). Along with system (1.9), we also consider the system with a constant forcing term in the right-hand side

$$x'(t) = (lx)(t) + f(t) + \mu, \quad t \in [a, b], \tag{5.1}$$

and the initial condition

$$x(a) = z, \tag{5.2}$$

where $\mu = \mathrm{col}(\mu_1, \dots, \mu_n)$ is a control parameter.

We will show that for arbitrary fixed $z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, and $\eta \in^{n-q}$, the parameter μ can be chosen so that the solution $x(\cdot, z, \lambda, \eta, \mu)$ of the initial value problem (5.1), (5.2) is, at the same time, a solution of the two-point parametrised boundary value problem (5.1), (2.3).

Proposition 5.1. *Let $z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, and $\eta \in \mathbb{R}^{n-q}$ be arbitrary given vectors. Assume that condition* (4.7) *is satisfied. Then a solution $x(\cdot)$ of the initial value problem* (5.1), (5.2) *satisfies the boundary conditions* (2.3) *if and only if $x(\cdot)$ coincides with $x_\infty(\cdot, z, \lambda, \eta)$ and*

$$\mu = \mu_{z,\lambda,\eta}, \tag{5.3}$$

where

$$\begin{aligned} \mu_{z,\lambda,\eta} := {}& \frac{1}{b-a}\left(D^{-1}(d - B\lambda + N_q\eta) - \left(D^{-1}A + \mathbb{1}_n\right)z\right) \\ & - \frac{1}{b-a}\int_a^b [(lx_\infty(\cdot, z, \lambda, \eta)(s) + f(s))]ds \end{aligned} \tag{5.4}$$

and $x_\infty(\cdot, z, \lambda, \eta)$ is the limit function of sequence (4.1).

Proof. The assertion of Proposition 5.1 is obtained by analogy to the proof of [50, Theorem 4.2]. Indeed, let $z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, and $\eta \in \mathbb{R}^{n-q}$ be arbitrary.

to Contents, page 1

If μ is given by (5.3), then, due to Theorem 4.2, the function $x_\infty(\cdot, z, \lambda, \eta)$ has properties (4.9) and satisfies equation (4.10), whence, by differentiation, equation (5.1) with the above-mentioned value of μ is obtained. Thus, $x_\infty(\cdot, z, \lambda, \eta)$ is a solution of (5.1), (5.2) with μ of form (5.3) and, moreover, this function satisfies the two-point boundary conditions (2.3).

Let us fix an arbitrary $\mu \in \mathbb{R}^n$ and assume that the initial value problem (5.1), (5.2) has a solution y satisfies the two-point boundary conditions (2.3). Then

$$y(t) = z + \int_a^t [(ly)(s) + f(s)] ds + \mu(t - a), \tag{5.5}$$

for all $t \in [a, b]$. By assumption, y satisfies the two-point conditions (2.3) and, therefore, (5.5) yields

$$\begin{aligned} Ay(a) + Dy(b) &= Az + D\left(z + \int_a^b ((ly)(s) + f(s))(s) ds + \mu(b - a) \right) \\ &= d - B\lambda + N_q \eta, \end{aligned} \tag{5.6}$$

whence we find that μ can be represented in the form

$$\mu = \frac{1}{b - a} D^{-1} \left(d - B\lambda + N_q \eta - (A + D)z - \int_a^b ((ly)(s) + f(s))(s) ds \right). \tag{5.7}$$

On the other hand, we already know that the function $x_\infty(\cdot, z, \lambda, \eta)$, satisfies the two-point conditions (2.3) and is a solution of the initial value problem (5.1), (5.2) with $\mu = \mu_{z,\lambda,\eta}$, where the value $\mu_{z,\lambda,\eta}$ is defined by formula (5.4). Consequently,

$$x_\infty(t, z, \lambda, \eta) = z + \int_a^t [(l x_\infty(\cdot, z, \lambda, \eta)(s) + f(s))] ds + \mu_{z,\lambda,\eta}(t - a), \quad t \in [a, b]. \tag{5.8}$$

Putting

$$h(t) := y(t) - x_\infty(t, z, \lambda, \eta), \quad t \in [a, b], \tag{5.9}$$

and taking (5.5), (5.8) into account, we obtain

$$h(t) = \int_a^t (lh)(s) ds + (\mu - \mu_{z,\lambda,\eta})(t - a), \quad t \in [a, b]. \tag{5.10}$$

to Contents, page 1

Recalling the definition (5.4) of $\mu_{z,\lambda,\eta}$ and using formula (5.7), we obtain

$$\begin{aligned}\mu - \mu_{z,\lambda,\eta} &= \frac{1}{b-a}\int_a^b l(x_\infty(\cdot,z,\lambda,\eta) - y)(s)ds \\ &= -\frac{1}{b-a}\int_a^b (lh)(s)ds,\end{aligned} \tag{5.11}$$

and, therefore, equality (5.10) can be rewritten as

$$\begin{aligned}h(t) &= \int_a^t (lh)(s)ds - \frac{t-a}{b-a}\int_a^b (lh)(s)ds \\ &= \left(1-\frac{t-a}{b-a}\right)\int_a^t (lh)(s)ds - \frac{t-a}{b-a}\int_t^b (lh)(s)ds, \quad t \in [a,b].\end{aligned} \tag{5.12}$$

Applying Lemma 3.2 and recalling notation (4.6), we get

$$\begin{aligned}|h(t)| &\le \left(\left(1-\frac{t-a}{b-a}\right)\int_a^t K_l(s)ds + \frac{t-a}{b-a}\int_t^b K_l(s)ds\right)\max_{\tau\in[a,b]}|h(\tau)| \\ &\le Q_l \max_{\tau\in[a,b]}|h(\tau)|\end{aligned} \tag{5.13}$$

for an arbitrary $t \in [a,b]$. By virtue of condition (4.7), inequality (5.13) implies that

$$\max_{\tau\in[a,b]}|h(\tau)| \le Q_l^m \max_{\tau\in[a,b]}|h(\tau)| \longrightarrow 0 \tag{5.14}$$

as $m \to +\infty$. According to (5.9), this means that y coincides with $x_\infty(\cdot,z,\lambda,\eta)$, and, therefore, by (5.11), $\mu = \mu_{z,\lambda,\eta}$, which brings us to the desired conclusion. □

We show that one can choose certain values of parameters $z = z_*$, $\lambda = \lambda_*$, $\eta = \eta_*$ for which the function $x_\infty(\cdot, z_*, \lambda_*, \eta_*)$ is the solution of the original three-point boundary value problem (1.9), (1.10). Let us consider the function $\Delta : \mathbb{R}^{3n-q} \to \mathbb{R}^n$ given by formula

$$\Delta(z,\lambda,\eta) := g(z,\lambda,\eta) - \int_a^b ((lx_\infty(\cdot,z,\lambda,\eta))(s) + f(s))ds \tag{5.15}$$

with

$$g(z,\lambda,\eta) := D^{-1}(d - B\lambda + N_q\eta) - \left(D^{-1}A + \mathbb{1}_n\right)z \tag{5.16}$$

for all z, λ, and η, where x_∞ is the limit function (4.8).

The following statement shows the relation of the limit function (4.8) to the solution of the original three-point boundary value problem (1.9), (1.10).

to Contents, page 1

Theorem 5.2. *Assume condition* (4.7). *Then the function* $x_\infty(\cdot, z, \lambda, \eta)$ *is a solution of the three-point boundary value problem* (1.9), (1.10) *if and only if the triplet* z, λ, η *satisfies the system of* $3n - q$ *algebraic equations*

$$\Delta(z, \lambda, \eta) = 0, \tag{5.17}$$

$$e_1 x_\infty(\ \xi, z, \lambda, \eta) = \lambda_1,\ e_2 x_\infty(\xi, z, \lambda, \eta) = \lambda_2,\ \ldots,\ e_n x_\infty(\xi, z, \lambda, \eta) = \lambda_n, \tag{5.18}$$

$$e_{q+1} x_\infty(b, z, \lambda, \eta) = \eta_1,\ e_{q+2} x_\infty(b, z, \lambda, \eta) = \eta_2, \ldots,\ e_{q+\infty} x_\infty(b, z, \lambda, \eta) = \eta_{n-q}. \tag{5.19}$$

Proof. It is sufficient to apply Proposition 5.1 and notice that the differential equation in (5.1) coincides with (1.9) if and only if the triplet (z, λ, η) satisfies (5.17). On the other hand, (5.18) and (5.19) bring us from the auxiliary two-point parametrised conditions to the three-point conditions (1.10). □

Proposition 5.3. *Assume condition* (4.7). *Then, for any* (z^j, λ^j, η^j), $j = 0, 1$, *the estimate*

$$\max_{t\in[a,b]} \left| x_\infty\left(t, z^0, \lambda^0, \eta^0\right) - x_\infty\left(t, z^1, \lambda^1, \eta^1\right) \right| \le (\mathbb{1}_n - Q_l)^{-1} v\left(z^0, \lambda^0, \eta^0, z^1, \lambda^1, \eta^1\right) \tag{5.20}$$

holds, where

$$v\left(z^0, \lambda^0, \eta^0, z^1, \lambda^1, \eta^1\right) := \max_{t\in[a,b]} \left| \varphi_{z^0,\lambda^0,\eta^0}(t) - \varphi_{z^1,\lambda^1,\eta^1}(t) \right|. \tag{5.21}$$

Proof. Let us fix two arbitrary triplets (z^j, λ^j, η^j), $j = 0, 1$, and put

$$u_m(t) := x_m\left(t, z^0, \lambda^0, \eta^0\right) - x_m\left(t, z^1, \lambda^1, \eta^1\right), \quad t \in [a, b]. \tag{5.22}$$

Consider the sequence of vectors c_m, $m = 0, 1, \ldots$, determined by the recurrence relation

$$c_m := c_0 + Q_l c_{m-1}, \quad m \ge 1, \tag{5.23}$$

with

$$c_0 := v\left(z^0, \lambda^0, \eta^0, z^1, \lambda^1, \eta^1\right). \tag{5.24}$$

Let us show that

$$\max_{t\in[a,b]} |u_m(t)| \le c_m \tag{5.25}$$

to Contents, page 1

for all $m \geq 0$. Indeed, estimate (5.25) is obvious for $m = 0$. Assume that

$$\max_{t\in[a,b]}|u_{m-1}(t)| \leq c_{m-1}. \tag{5.26}$$

It follows immediately from (4.1) that

$$\begin{aligned} u_m(t) &= \varphi_{z^0,\lambda^0,\eta^0}(t) - \varphi_{z^1,\lambda^1,\eta^1}(t) + \int_a^t (lu_{m-1})(s)ds - \frac{t-a}{b-a}\int_a^b (lu_{m-1})(s)ds \\ &= \varphi_{z^0,\lambda^0,\eta^0}(t) - \varphi_{z^1,\lambda^1,\eta^1}(t) \\ &\quad + \left(1 - \frac{t-a}{b-a}\right)\int_a^t (lu_{m-1})(s)ds - \frac{t-a}{b-a}\int_t^b (lu_{m-1})(s)ds, \end{aligned} \tag{5.27}$$

whence, by virtue of (5.21), estimate (3.7) to Lemma 3.2, and assumption (5.26),

$$\begin{aligned} |u_m(t)| &\leq \left|\varphi_{z^0,\lambda^0,\eta^0}(t) - \varphi_{z^1,\lambda^1,\eta^1}(t)\right| \\ &\quad + \left(1 - \frac{t-a}{b-a}\right)\int_a^t |(lu_{m-1})(s)|ds + \frac{t-a}{b-a}\int_t^b |(lu_{m-1})(s)|ds \\ &\leq v\left(z^0,\lambda^0,\eta^0,z^1,\lambda^1,\eta^1\right) \\ &\quad + \left(1 - \frac{t-a}{b-a}\right)\int_a^t K_l(s)ds \max_{t\in[a,b]}|u_{m-1}(t)| + \frac{t-a}{b-a}\int_t^b K_l(s)ds \max_{t\in[a,b]}|u_{m-1}(t)| \\ &\leq v\left(z^0,\lambda^0,\eta^0,z^1,\lambda^1,\eta^1\right) + \left(\left(1 - \frac{t-a}{b-a}\right)\int_a^t K_l(s)ds + \frac{t-a}{b-a}\int_t^b K_l(s)ds\right)c_{m-1} \\ &\leq v\left(z^0,\lambda^0,\eta^0,z^1,\lambda^1,\eta^1\right) + Q_l c_{m-1}, \end{aligned} \tag{5.28}$$

which estimate, in view of (5.23) and (5.24), coincides with the required inequality (5.25). Thus, (5.25) is true for any m. Using (5.23) and (5.25), we obtain

$$\begin{aligned} \max_{t\in[a,b]}|u_m(t)| &\leq c_0 + Q_l c_{m-1} = c_0 + Q_l c_0 + Q_l^2 c_{m-2} = \cdots \\ &= \sum_{k=0}^{m-1} Q_l^k c_0 + Q_l^m c_0. \end{aligned} \tag{5.29}$$

Due to assumption (4.7), $\lim_{m\to+\infty} Q_l^m = 0$. Therefore, passing to the limit in (5.29) as $m \to +\infty$ and recalling notation (5.22), we obtain the estimate

$$\max_{t\in[a,b]}\left|x_*\left(t,z^0,\lambda^0,\eta^0\right) - x_*\left(t,z^1,\lambda^1,\eta^1\right)\right| \leq \sum_{k=0}^{+\infty} Q_l^k c_0 = (\mathbb{1}_n - Q_l)^{-1} c_0, \tag{5.30}$$

to Contents, page 1

which, in view of (5.24), coincides with (5.20). □

Now we establish some properties of the "determining function" $\Delta : \mathbb{R}^{3n-q} \to \mathbb{R}^n$ given by equality (5.15).

Proposition 5.4. *Under condition* (3.10)*, formula* (5.15) *determines a well-defined function* $\Delta : \mathbb{R}^{3n-q} \to \mathbb{R}^n$*, which, moreover, satisfies the estimate*

$$\left|\Delta\left(z^0,\lambda^0,\eta^0\right)-\Delta\left(z^1,\lambda^1,\eta^1\right)\right| \le \left|G[z^0-z^1,\lambda^0-\lambda^1,\eta^0-\eta^1]^*\right| + R_l \max_{t\in[a,b]}\left|z^0-z^1+\frac{t-a}{b-a}G[z^0-z^1,\lambda^0-\lambda^1,\eta^0-\eta^1]^*\right|, \tag{5.31}$$

for all (z^j,λ^j,η^j)*,* $j=0,1$*, where the* $(n\times n)$*-matrices* G *and* R_l *are defined by the equalities*

$$G := D^{-1}\left[A+D,B,N_q\right], \qquad R_l := \int_a^b K_l(s)ds\ (\mathbb{1}_n - Q_l)^{-1}. \tag{5.32}$$

Proof. According to the definition (5.15) of Δ, we have

$$\Delta\left(z^0,\lambda^0,\eta^0\right)-\Delta\left(z^1,\lambda^1,\eta^1\right) = g\left(z^0,\lambda^0,\eta^0\right)-g\left(z^1,\lambda^1,\eta^1\right) - \int_a^b\left(l\left(x_\infty\left(\cdot,z^0,\lambda^0,\eta^0\right)-x_\infty\left(\cdot,z^1,\lambda^1,\eta^1\right)\right)(s)\right)ds, \tag{5.33}$$

whence, due to Lemma 3.2,

$$\begin{aligned}\left|\Delta\left(z^0,\lambda^0,\eta^0\right)-\Delta\left(z^1,\lambda^1,\eta^1\right)\right| &\le \left|g\left(z^0,\lambda^0,\eta^0\right)-g\left(z^1,\lambda^1,\eta^1\right)\right| \\ &\quad+\int_a^b\left|l\left(x_\infty\left(\cdot,z^0,\lambda^0,\eta^0\right)-x_\infty\left(\cdot,z^1,\lambda^1,\eta^1\right)\right)(s)\right|ds \\ &\le \left|g\left(z^0,\lambda^0,\eta^0\right)-g\left(z^1,\lambda^1,\eta^1\right)\right| \\ &\quad+\int_a^b K_l(s)ds\max_{\tau\in[a,b]}\left|x_\infty\left(\tau,z^0,\lambda^0,\eta^0\right)-x_\infty\left(\tau,z^1,\lambda^1,\eta^1\right)(s)\right|.\end{aligned} \tag{5.34}$$

to Contents, page 1

Using Proposition 5.3, we find

$$\begin{aligned}\left|\Delta\left(z^0,\lambda^0,\eta^0\right)-\Delta\left(z^1,\lambda^1,\eta^1\right)\right| \le{}& \left|g\left(z^0,\lambda^0,\eta^0\right)-g\left(z^1,\lambda^1,\eta^1\right)\right|\\ &+\int_a^b K_l(s)ds(\mathbb{1}_n-Q_l)^{-1}v\left(z^0,\lambda^0,\eta^0,z^1,\lambda^1,\eta^1\right).\end{aligned} \tag{5.35}$$

On the other hand, recalling (4.2) and (5.21), we get

$$v\left(z^0,\lambda^0,\eta^0,z^1,\lambda^1,\eta^1\right)=\max_{t\in[a,b]}\left|z^0-z^1+\frac{t-a}{b-a}\left(g\left(z^0,\lambda^0,\eta^0\right)-g\left(z^1,\lambda^1,\eta^1\right)\right)\right|. \tag{5.36}$$

It follows immediately from (5.16) that

$$\begin{aligned}g\left(z^0,\lambda^0,\eta^0\right)-g\left(z^1,\lambda^1,\eta^1\right)&=-D^{-1}B\left(\lambda^0-\lambda^1\right)-D^{-1}N_q\left(\eta^0-\eta^1\right)-\left(D^{-1}A+\mathbb{1}_n\right)\left(z^0-z^1\right)\\ &=-D^{-1}\left[B\left(\lambda^0-\lambda^1\right)+N_q\left(\eta^0-\eta^1\right)+(A+D)\left(z^0-z^1\right)\right]\\ &=D^{-1}[A+D,B,N_q]\begin{pmatrix}z^0-z^1\\ \lambda^0-\lambda^1\\ \eta^0-\eta^1\end{pmatrix}.\end{aligned} \tag{5.37}$$

Therefore, (5.35) and (5.36) yield the estimate

$$\begin{aligned}&\left|\Delta\left(z^0,\lambda^0,\eta^0\right)-\Delta\left(z^1,\lambda^1,\eta^1\right)\right|\\ &\quad\le\left|D^{-1}[A+D,B,N_q]\begin{pmatrix}z^0-z^1\\ \lambda^0-\lambda^1\\ \eta^0-\eta^1\end{pmatrix}\right|\\ &\qquad+\int_a^b K_l(s)ds(\mathbb{1}_n-Q_l)^{-1}\max_{t\in[a,b]}\left|z^0-z^1+\frac{t-a}{b-a}D^{-1}[A+D,B,N_q]\begin{pmatrix}z^0-z^1\\ \lambda^0-\lambda^1\\ \eta^0-\eta^1\end{pmatrix}\right|,\end{aligned} \tag{5.38}$$

which, in view of (5.32), coincides with (5.31). □

Properties stated by Propositions 5.3 and 5.4 can be used when analysing conditions guaranteeing the solvability of the determining equations.

to Contents, page 1

6. On the Numerical-Analytic Algorithm of Solving the Problem

Theorems 4.2 and 5.2 allow one to formulate the following numerical-analytic algorithm for the construction of a solution of the three-point boundary value problem (1.9), (1.10).

(1) For any vector $z \in \mathbb{R}^n$, according to (4.1), we analytically construct the sequence of functions $x_m(\cdot, z, \lambda, \eta)$ depending on the parameters z, λ, η and satisfying the auxiliary two-point boundary condition (2.3).

(2) We find the limit $x_\infty(\cdot, z, \lambda, \eta)$ of the sequence $x_m(\cdot, z, \lambda, \eta)$ satisfying (2.3).

(3) We construct the algebraic determining system (5.17), (5.18), and (5.19) with respect $3n - q$ scalar variables.

(4) Using a suitable numerical method, we (approximately) find a root

$$z_* \in \mathbb{R}^n, \qquad \lambda_* \in \mathbb{R}^n, \qquad \eta_* \in \mathbb{R}^{n-q} \tag{6.1}$$

of the determining system (5.17), (5.18), and (5.19).

(5) Substituting values (6.1) into $x_\infty(\cdot, z, \lambda, \eta)$, we obtain a solution of the original three-point boundary value problem (1.9), (1.10) in the form

$$x(t) = x_\infty(t, z_*, \lambda_*, \eta_*), \quad t \in [a, b]. \tag{6.2}$$

This solution (6.2) can also be obtained by solving the Cauchy problem

$$x(a) = z_* \tag{6.3}$$

for (1.9).

The fundamental difficulty in the realization of this approach arises at point (2) and is related to the analytic construction of the function $x_\infty(\cdot, z, \lambda, \eta)$. This problem can often be overcome by considering certain approximations of form (4.1), which, unlike the function $x_\infty(\cdot, z, \lambda, \eta)$, are known in the analytic form. In practice, this means that we fix a suitable $m \geq 1$, construct the corresponding function $x_m(\cdot, z, \lambda, \eta)$ according to relation (4.1), and define the function $\Delta_m : \mathbb{R}^{3n-q} \to \mathbb{R}^n$ by putting

$$\Delta_m(z, \lambda, \eta) := D^{-1}(d - B\lambda + N_q\eta) - \left(D^{-1}A + \mathbb{1}_n\right)z - \int_a^b \left[(lx_m(\cdot, z, \lambda, \eta)(s) + f(s))\right] ds, \tag{6.4}$$

for arbitrary z, λ, and η. To investigate the solvability of the three-point boundary value problem (1.9), (1.10), along with the determining system (5.17), (5.18), and (5.19), one considers the mth approximate determining system

$$\begin{gathered}\Delta_m(z, \lambda, \eta) = 0, \\ e_1 x_m(\xi, z, \lambda, \eta) = \lambda_1, e_2 x_m(\xi, z, \lambda, \eta) = \lambda_2, \dots, e_n x_m(\xi, z, \lambda, \eta) = \lambda_n, \\ e_{q+1} x_m(b, z, \lambda, \eta) = \eta_1, \dots, e_n x_m(b, z, \lambda, \eta) = \eta_{n-q},\end{gathered} \tag{6.5}$$

to Contents, page 1

where e_i, $i = 1, 2, \ldots, n$, are the vectors given by (5.15).

It is natural to expect (and, in fact, can be proved) that, under suitable conditions, the systems (5.17), (5.18), (5.19), and (6.5) are "close enough" to one another for m sufficiently large. Based on this circumstance, existence theorems for the three-point boundary value problem (1.9), (1.10) can be obtained by studying the solvability of the approximate determining system (6.5) (in the case of periodic boundary conditions, see, e.g., [35]).

Acknowledgments

This research was carried out as part of the TAMOP-4.2.1.B-10/2/KONV-2010-0001 project with support by the European Union, co-financed by the European Social Fund.

References

[1] R. P. Agarwal and D. O'Regan, "A survey of recent results for initial and boundary value problems singular in the dependent variable," in *Handbook of Differential Equations*, pp. 1–68, Elsevier, Amsterdam, The Netherlands, 2004.

[2] N. N. Bogoljubov, J. A. Mitropol'skiĭ, and A. M. Samoĭlenko, *The Method of Accelerated Convergence in Nonlinear Mechanics*, Naukova Dumka, Kiev, Ukraine, 1969.

[3] M. Farkas, *Periodic Motions*, vol. 104 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 1994.

[4] M. Fečkan, "Minimal periods of periodic solutions," *Miskolc Mathematical Notes*, vol. 7, no. 2, pp. 121–139, 2006.

[5] I. Kiguradze and S. Mukhigulashvili, "On periodic solutions of two-dimensional nonautonomous differential systems," *Nonlinear Analysis*, vol. 60, no. 2, pp. 241–256, 2005.

[6] I. T. Kiguradze, *Some Singular Boundary Value Problems for Ordinary Differential Equations*, Izdat. Tbilis. Univ., Tbilisi, Georgia, 1975.

[7] E. F. Mishchenko and N. K. Rozov, *Differential Equations with Small Parameters and Relaxation Oscillations*, vol. 13 of *Mathematical Concepts and Methods in Science and Engineering*, Plenum Press, New York, NY, USA, 1980, translated from the Russian by F. M. C. Goodspee.

[8] J. Moser, "A rapidly convergent iteration method and non-linear differential equations. II," *Annali della Scuola Normale Superiore di Pisa*, vol. 20, pp. 499–535, 1966.

[9] J. Moser, "A rapidly convergent iteration method and non-linear partial differential equations. I," *Annali della Scuola Normale Superiore di Pisa*, vol. 20, pp. 265–315, 1966.

[10] S. K. Ntouyas, "Nonlocal initial and boundary value problems: a survey," in *Handbook of Differential Equations: Ordinary Differential Equations. Vol. II*, pp. 461–557, Elsevier, Amsterdam, The Netherlands, 2005.

[11] F. Sadyrbaev, "Multiplicity of solutions for second order two-point boundary value problems with asymptotically asymmetric nonlinearities at resonance," *Georgian Mathematical Journal*, vol. 14, no. 2, pp. 351–360, 2007.

[12] M. Ronto and A. M. Samoilenko, *Numerical-Analytic Methods in the Theory of Boundary-Value Problems*, World Scientific, River Edge, NJ, USA, 2000, with a preface by Yu. A. Mitropolsky and an appendix by the authors and S. I. Trofimchu.

[13] R. P. Agarwal, D. O'Regan, and S. Staněk, "Positive and maximal positive solutions of singular mixed boundary value problem," *Central European Journal of Mathematics*, vol. 7, no. 4, pp. 694–716, 2009.

[14] A. Cabada, A. Lomtatidze, and M. Tvrdý, "Periodic problem involving quasilinear differential operator and weak singularity," *Advanced Nonlinear Studies*, vol. 7, no. 4, pp. 629–649, 2007.

[15] A. Capietto, J. Mawhin, and F. Zanolin, "A continuation approach to superlinear periodic boundary value problems," *Journal of Differential Equations*, vol. 88, no. 2, pp. 347–395, 1990.

[16] A. Capietto, J. Mawhin, and F. Zanolin, "Continuation theorems for periodic perturbations of autonomous systems," *Transactions of the American Mathematical Society*, vol. 329, no. 1, pp. 41–72, 1992.

to Contents, page 1

[17] L. Cesari, "Functional analysis and periodic solutions of nonlinear differential equations," *Contributions to Differential Equations*, vol. 1, pp. 149–187, 1963.
[18] L. Cesari, *Nonlinear Analysis and Alternative Methods*, Accademia Nazionale dei Lincei, Rome, Italy, 1977, lecture series held in April 1974, contributi del centro linceo interdisciplinare di scienze matematiche e loro applicazioni, 39.
[19] R. E. Gaines and J. L. Mawhin, *Coincidence Degree, and Nonlinear Differential Equations*, vol. 568 of *Lecture Notes in Mathematics*, Springer, Berlin, Germany, 1977.
[20] J. K. Hale, *Oscillations in Nonlinear Systems*, McGraw-Hill, New York, NY, USA, 1963.
[21] M. A. Krasnosel'skiĭ, G. M. Vaĭnikko, P. P. Zabreĭko, Ya. B. Rutitskii, and V. Y. Stetsenko, *Approximate Solution of Operator Equations*, Wolters-Noordhoff, Groningen, The Netherlands, 1972, translated from the Russian by D. Louvis.
[22] M. A. Krasnosel'skiĭ and P. P. Zabreĭko, *Geometrical Methods of Nonlinear Analysis*, vol. 263 of *Fundamental Principles of Mathematical Sciences*, Springer, Berlin, Germany, 1984, translated from the Russian by C. C. Fenske.
[23] J. Mawhin, *Topological Degree Methods in Nonlinear Boundary Value Problems*, vol. 40 of *CBMS Regional Conference Series in Mathematics*, American Mathematical Society, Providence, RI, USA, 1979, expository lectures from the CBMS Regional Conference held at Harvey Mudd College, Claremont, Calif., June, 1977.
[24] I. Rachunková, "Strong singularities in mixed boundary value problems," *Mathematica Bohemica*, vol. 131, no. 4, pp. 393–409, 2006.
[25] I. Rachunková, S. Staněk, and M. Tvrdý, *Solvability of Nonlinear Singular Problems for Ordinary Differential Equations*, vol. 5 of *Contemporary Mathematics and Its Applications*, Hindawi Publishing Corporation, New York, NY, USA, 2008.
[26] S. Staněk, "Positive solutions of singular Dirichlet boundary value problems with time and space singularities," *Nonlinear Analysis*, vol. 71, no. 10, pp. 4893–4905, 2009.
[27] U. M. Ascher, R. M. M. Mattheij, and R. D. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, vol. 13 of *Classics in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 1995, corrected reprint of the 1988 original.
[28] H. B. Keller, *Numerical Methods for Two-Point Boundary Value Problems*, Dover, New York, NY, USA, 1992, corrected reprint of the 1968 edition.
[29] L. Cesari, *Asymptotic Behavior and Stability Problems in Ordinary Differential Equations*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 16, Springer, New York, NY, USA, 3rd edition, 1971.
[30] A. M. Samoilenko, "A numerical-analytic method for investigation of periodic systems of ordinary differential equations. I," *Ukrainskiĭ Matematicheskiĭ Zhurnal*, vol. 17, no. 4, pp. 82–93, 1965.
[31] A. M. Samoilenko, "A numerical-analytic method for investigation of periodic systems of ordinary differential equations. II," *Ukrainskiĭ Matematicheskiĭ Zhurnal*, vol. 18, no. 2, pp. 50–59, 1966.
[32] A. M. Samoilenko and N. I. Ronto, *Numerical-Analytic Methods of Investigating Periodic Solutions*, Mir, Moscow, Russia, 1979, translated from the Russian by V. Shokurov, with a foreword by Yu. A. Mitropolskii.
[33] A. M. Samoilenko and N. I. Ronto, *Numerical-Analytic Methods of Investigation of Boundary-Value Problems*, Naukova Dumka, Kiev, Ukraine, 1986, with an English summary, edited and with a preface by Yu. A. Mitropolskii.
[34] A. M. Samoilenko and N. I. Ronto, *Numerical-Analytic Methods in the Theory of Boundary-Value Problems for Ordinary Differential Equations*, Naukova Dumka, Kiev, Ukraine, 1992, edited and with a preface by Yu. A. Mitropolskii.
[35] A. Rontó and M. Rontó, "Successive approximation techniques in non-linear boundary value problems for ordinary differential equations," in *Handbook of Differential Equations: Ordinary Differential Equations. Vol. IV*, Handbook of Differential Equations, pp. 441–592, Elsevier, Amsterdam, The Netherlands, 2008.
[36] A. Augustynowicz and M. Kwapisz, "On a numerical-analytic method of solving of boundary value problem for functional-differential equation of neutral type," *Mathematische Nachrichten*, vol. 145, pp. 255–269, 1990.
[37] N. A. Evkhuta and P. P. Zabreiko, "A. M. Samoilenko's method for finding periodic solutions of quasilinear differential equations in a Banach space," *Ukrainian Mathematical Journal*, vol. 37, no. 2, p. 162–168, 269, 1985.
[38] T. Jankowski, "Numerical-analytic method for implicit differential equations," *Miskolc Mathematical Notes*, vol. 2, no. 2, pp. 137–144, 2001.

to Contents, page 1

[39] T. Jankowski, "The application of numerical-analytic method for systems of differential equations with a parameter," *Ukrainian Mathematical Journal*, vol. 54, no. 4, pp. 545–554, 2002.

[40] T. Jankowski, "Numerical-analytic methods for differential-algebraic systems," *Acta Mathematica Hungarica*, vol. 95, no. 3, pp. 243–252, 2002.

[41] T. Jankowski, "Monotone and numerical-analytic methods for differential equations," *Computers & Mathematics with Applications*, vol. 45, no. 12, pp. 1823–1828, 2003.

[42] M. Kwapisz, "On modifications of the integral equation of Samoĭlenko's numerical-analytic method of solving boundary value problems," *Mathematische Nachrichten*, vol. 157, pp. 125–135, 1992.

[43] M. Kwapisz, "Some remarks on an integral equation arising in applications of numerical-analytic method of solving of boundary value problems," *Ukrainian Mathematical Journal*, vol. 44, no. 1, pp. 115–119, 1992.

[44] A. Perov, L. Dikareva, S. Oleinikova, and M. Portnov, "On the convergence analysis of A. M. Samoilenko's method," *Vestnik Voronezhskogo Gosudarstvennogo Universiteta. Seriya Fizika, Matematika*, pp. 111–119, 2001.

[45] A. Ronto and M. Rontó, "A note on the numerical-analytic method for nonlinear two-point boundary-value problems," *Nonlinear Oscillations*, vol. 4, no. 1, pp. 112–128, 2001.

[46] A. Rontó and M. Rontó, "On a Cauchy-Nicoletti type three-point boundary value problem for linear differential equations with argument deviations," *Miskolc Mathematical Notes*, vol. 10, no. 2, pp. 173–205, 2009.

[47] A. Rontó and M. Rontó, "Successive approximation method for some linear boundary value problems for differential equations with a special type of argument deviation," *Miskolc Mathematical Notes*, vol. 10, no. 1, pp. 69–95, 2009.

[48] A. N. Ronto, "On some boundary value problems for Lipschitz differential equations," *Nelīnīĭnī Kolivannya*, pp. 74–94, 1998.

[49] A. N. Ronto, M. Rontó, A. M. Samoilenko, and S. I. Trofimchuk, "On periodic solutions of autonomous difference equations," *Georgian Mathematical Journal*, vol. 8, no. 1, pp. 135–164, 2001.

[50] A. N. Ronto, M. Ronto, and N. M. Shchobak, "On the parametrization of three-point nonlinear boundary value problems," *Nonlinear Oscillations*, vol. 7, no. 3, pp. 384–402, 2004.

[51] N. I. Ronto, A. M. Samoilenko, and S. I. Trofimchuk, "The theory of the numerical-analytic method: achievements and new directions of development. I," *Ukrainian Mathematical Journal*, vol. 50, no. 1, pp. 116–135, 1998.

[52] N. I. Ronto, A. M. Samoilenko, and S. I. Trofimchuk, "The theory of the numerical-analytic method: achievements and new directions of development. II," *Ukrainian Mathematical Journal*, vol. 50, no. 2, pp. 255–277, 1998.

[53] N. I. Ronto, A. M. Samoilenko, and S. I. Trofimchuk, "The theory of the numerical-analytic method: achievements and new directions of development. III," *Ukrainian Mathematical Journal*, vol. 50, no. 7, pp. 1091–1114, 1998.

[54] N. I. Ronto, A. M. Samoilenko, and S. I. Trofimchuk, "The theory of the numerical-analytic method: achievements and new directions of development. IV," *Ukrainian Mathematical Journal*, vol. 50, no. 12, pp. 1888–1907, 1999.

[55] N. I. Ronto, A. M. Samoilenko, and S. I. Trofimchuk, "The theory of the numerical-analytic method: achievements and new directions of development. V," *Ukrainian Mathematical Journal*, vol. 51, no. 5, pp. 735–747, 2000.

[56] N. I. Ronto, A. M. Samoilenko, and S. I. Trofimchuk, "The theory of the numerical-analytic method: achievements and new directions of development. VI," *Ukrainian Mathematical Journal*, vol. 51, no. 7, pp. 1079–1094, 2000.

[57] N. I. Ronto, A. M. Samoilenko, and S. I. Trofimchuk, "The theory of the numerical-analytic method: achievements and new directions of development. VII," *Ukrainian Mathematical Journal*, vol. 51, no. 9, pp. 1399–1418, 2000.

[58] M. Rontó and J. Mészáros, "Some remarks on the convergence of the numerical-analytical method of successive approximations," *Ukrainian Mathematical Journal*, vol. 48, no. 1, pp. 101–107, 1996.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 134072, 9 pages
doi:10.1155/2011/134072

Research Article

On Stability of Linear Delay Differential Equations under Perron's Condition

J. Diblík[1, 2] and A. Zafer[3]

[1] *Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Brno University of Technology, Brno 602 00, Czech Republic*
[2] *Department of Mathematics, Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno 616 00, Czech Republic*
[3] *Department of Mathematics, Middle East Technical University, 06531 Ankara, Turkey*

Correspondence should be addressed to J. Diblík, diblik@feec.vutbr.cz

Received 18 January 2011; Accepted 22 February 2011

Academic Editor: Miroslava Růžičková

The stability of the zero solution of a system of first-order linear functional differential equations with nonconstant delay is considered. Sufficient conditions for stability, uniform stability, asymptotic stability, and uniform asymptotic stability are established.

1. Introduction

We begin with a classical result for the linear system

$$x' = A(t)x, \tag{L1}$$

where A is an $n \times n$ matrix function defined and continuous on $[0, \infty)$. By $C_B[0, \infty)$, we will denote the set of bounded functions defined and continuous on $[0, \infty)$ and by $|\cdot|$ the Euclidean norm.

In 1930, Perron first formulated the following definition being named after him.

Definition 1.1 (see [1]). System (L1) is said to satisfy Perron's condition (P1) if, for any given vector function $f \in C_B[0, \infty)$, the solution $x(t)$ of

$$x' = A(t)x + f(t), \quad x(0) = 0 \tag{N1}$$

is bounded.

to Contents, page 1

The following theorem by Bellman [2] is well known.

Theorem 1.2 (see [2]). *If* (P1) *holds and* $|A(t)| \leq M_1$ *for some positive number* M_1*, then the zero solution of* (L1) *is uniformly asymptotically stable.*

The proof is accomplished by making use of the basic properties of a fundamental matrix, the Banach-Steinhaus theorem, and the adjoint system

$$x' = -A^T(t)x, \tag{1.1}$$

where A^T denotes the transpose of A.

It is shown by an example in [3] that Theorem 1.2 may not be valid if the function f appearing in (N1) is replaced by a constant vector. However, such a theorem is later obtained in [4] under a *Perron-like* condition.

Theorem 1.2 is extended by Halanay [5] to linear delay systems of the form

$$x'(t) = A(t)x(t) + B(t)x(t-\tau), \tag{L2}$$

where A, B are $n \times n$ matrix functions defined and continuous on $[0, \infty)$ and τ is a positive real number.

Definition 1.3. System (L2) is said to satisfy Perron's condition (P2) if for any given vector function $f \in C_B[0, \infty)$, the solution $x(t)$ of

$$x'(t) = A(t)x(t) + B(t)x(t-\tau) + f(t) \tag{N2}$$

satisfying $x(t) = 0$, $t \leq 0$, is bounded.

Theorem 1.4 (see [5]). *If* (P2) *holds,* $|A(t)| \leq M_1$*, and* $|B(t)| \leq M_2$ *for some positive numbers* M_1 *and* M_2*, then the zero solution of* (L2) *is uniformly asymptotically stable.*

The method used to prove Theorem 1.4 is similar to Bellman's except that the adjoint system

$$y'(t) = -A^T(t)y(t) - B^T(t+\tau)y(t+\tau) \tag{1.2}$$

is not constructed with respect to an inner product but the functional

$$F(x, y)(t) = \int_t^{t+\tau} y^T(s)B(s)x(s-\tau)ds + x^T(t)y(t). \tag{1.3}$$

For some extensions to impulsive differential equations, we refer the reader in particular to [6, 7].

In this paper, we consider the more general linear delay system

$$x'(t) = A(t)x(t) + B(t)x(g(t)), \tag{1.4}$$

to Contents, page 1

where A and B are $n \times n$ matrix functions defined and continuous on $[0, \infty)$ and g is a continuously differentiable increasing function defined on $[0, \infty)$ satisfying $g(t) < t$ and $g'(t) \leq 1$. We set $h := g^{-1}$. Obviously, $h \in C^1[0, \infty)$ and increases on $[0, \infty)$ and $h(t) > t$.

Perron's condition takes the following form.

Definition 1.5. System (1.4) is said to satisfy Perron's condition (P) if, for any given vector function $f \in C_B[0, \infty)$, the solution $x(t)$ of

$$x'(t) = A(t)x(t) + B(t)x\big(g(t)\big) + f(t) \tag{1.5}$$

satisfying $x(t) = 0$, $t \leq 0$ is bounded.

A natural question is whether the zero solution of (1.4) is uniformly asymptotically stable under Perron's condition (P). It turns out that the answer depends on the delay function g.

The paper is organized as follows. In Section 2, we only state our results; the proofs are included in Section 5. We define an adjoint system and give a variation of parameters formula in Section 3 to be needed in proving the main results. Section 4 contains also some lemmas concerning Perron's condition and a relation useful for changing the order of integration.

2. Stability Theorems

The conclusion obtained by Bellman and Halanay for systems (L1) and (L2), respectively, is quite strong. We are only able to prove the stability of the zero solution for more general equation (1.4) under Perron's condition. To get uniform stability or asymptotic stability or uniform asymptotic stability, we impose restrictions on the delay function.

For our purpose, we denote

$$\begin{aligned} h_*(t) &:= h(t) - t, \quad t \geq 0, \\ g_*(t, t_0) &:= \sup_{r \in [h(t_0), t]} \{r - g(r)\}, \quad t, t_0 \geq 0. \end{aligned} \tag{2.1}$$

Theorem 2.1. *Let* (P) *hold. If there are positive numbers* M_1 *and* M_2 *such that*

$$|A(t)| \leq M_1, \quad |B(t)| \leq M_2 \quad \forall t \geq 0, \tag{2.2}$$

then the zero solution of (1.4) *is stable.*

Theorem 2.2. *Let* (P) *hold. If* (2.2) *is satisfied and if there exists a positive real number* M_3 *such that*

$$h_*(t) \leq M_3 \quad \forall t \geq 0, \tag{2.3}$$

then the zero solution of (1.4) *is uniformly stable.*

to Contents, page 1

Theorem 2.3. *Let* (P) *hold. If* (2.2) *and*

$$\limsup_{t\to\infty}\frac{g_*(t,t_0)}{t-t_0}=0 \quad \text{for each } t_0\ge 0 \tag{2.4}$$

are satisfied, then the zero solution of (1.4) *is asymptotically stable.*

Theorem 2.4. *Let* (P) *hold. If* (2.2), (2.3), *and*

$$\limsup_{t\to\infty}\frac{g_*(t,t_0)}{t-t_0}=0 \quad \text{uniformly for } t_0\ge 0 \tag{2.5}$$

are satisfied, then the zero solution of (1.4) *is uniformly asymptotically stable.*

Remark 2.5. Note that if $g(t)=t-\tau$, then $h(t)=t+\tau$ and hence the conditions (2.3), (2.4), and (2.5) are automatically satisfied. In this case, all theorems become equivalent, that is, the zero solution is uniformly asymptotically stable. Thus, the results obtained by Bellman and Halanay are recovered.

3. Variation of Parameters Formula

To establish a variation of parameters formula to represent the solutions of (1.5), one needs an adjoint system. The following lemma helps to define the adjoint of (1.4).

Lemma 3.1. *Let* $x(t)$ *be a solution of* (1.4). *If* $y(t)$ *is a solution of*

$$y'(t)=-A^T(t)y(t)-B^T(h(t))y(h(t))h'(t), \tag{3.1}$$

then

$$\frac{d}{dt}F(x(t),y(t))=0, \tag{3.2}$$

where

$$F(x,y)(t)=\int_t^{h(t)}y^T(s)B(s)x(g(s))ds+x^T(t)y(t). \tag{3.3}$$

Proof. Verify directly. □

Definition 3.2. The system (3.1) is said to be adjoint to system (1.4).

It is easy to see that the adjoint of system (3.1) is system (1.4); thus the systems are mutually adjoint to each other.

to Contents, page 1

Lemma 3.3. *Let $Y(t,s)$ be a matrix solution of* (3.1) *for $t < s$ satisfying $Y(s,s) = I$ and $Y(t,s) = 0$ for $t > s$. Then $x(t)$ is a solution of* (1.5) *if and only if*

$$x(t) = Y^T(s,t)x(s) + \int_{g(s)}^{s} Y^T(h(\beta),t)B(h(\beta))x(\beta)h'(\beta)d\beta + \int_{s}^{t} Y^T(\beta,t)f(\beta)d\beta. \quad (3.4)$$

Proof. Replacing t by β in (1.5) and then integrating the resulting equation multiplied by $Y^T(\beta,t)$ over $\beta \in [s,t]$, we have

$$\begin{aligned}
&\int_s^t Y^T(\beta,t)A(\beta)x(\beta)d\beta + \int_s^t Y^T(\beta,t)B(\beta)x(g(\beta))d\beta + \int_s^t Y^T(\beta,t)f(\beta)d\beta \\
&\quad = \int_s^t Y^T(\beta,t)x'(\beta)d\beta \\
&\quad = x(t) - Y^T(s,t)x(s) - \int_s^t \left[\frac{\partial}{\partial\beta}Y^T(\beta,t)\right]x(\beta)d\beta \\
&\quad = x(t) - Y^T(s,t)x(s) + \int_s^t \left[Y^T(\beta,t)A(\beta) + Y^T(h(\beta),t)B(h(\beta))h'(\beta)\right]x(\beta)d\beta \\
&\quad = x(t) - Y^T(s,t)x(s) + \int_s^t Y^T(\beta,t)A(\beta)x(\beta)d\beta + \int_{h(s)}^{h(t)} Y^T(\beta,t)B(\beta)x(g(\beta))d\beta.
\end{aligned} \quad (3.5)$$

Comparing both sides and using

$$\int_t^{h(t)} Y^T(\beta,t)B(\beta)x(g(\beta))d\beta = 0, \quad (3.6)$$

which is true in view of $Y(\beta,t) = 0$ for $\beta > t$, we get

$$x(t) = Y^T(s,t)x(s) - \int_{h(s)}^{s} Y^T(\beta,t)B(\beta)x(g(\beta))d\beta + \int_s^t Y^T(\beta,t)f(\beta)d\beta \quad (3.7)$$

and hence

$$x(t) = Y^T(s,t)x(s) + \int_{g(s)}^{s} Y^T(h(\beta),t)B(h(\beta))x(\beta)h'(\beta)d\beta + \int_{s}^{t} Y^T(\beta,t)f(\beta)d\beta. \quad (3.8)$$

□

It is not difficult to see from (3.4) that if $X(t,s)$ is a matrix solution of (1.4) for $t > s$ satisfying $X(s,s) = I$ and $X(t,s) = 0$ for $t < s$, then

$$X(t,s) = Y^T(s,t). \quad (3.9)$$

Using this relation in Lemma 3.3 leads to the following variation of parameters formula.

to Contents, page 1

Lemma 3.4. *Let $X(t,s)$ be a matrix solution of* (1.4) *for $t > s$ satisfying $X(s,s) = I$ and $X(t,s) = 0$ for $t < s$. Then $x(t)$ is a solution of* (1.5) *if and only if*

$$x(t) = X(t,s)x(s) + \int_{g(s)}^{s} X(t,h(\beta))B(h(\beta))x(\beta)h'(\beta)d\beta + \int_{s}^{t} X(t,\beta)f(\beta)d\beta. \tag{3.10}$$

4. Auxiliary Results

Lemma 4.1. *If* (P) *holds, then there is a positive number K_1 such that*

$$\int_0^t |X(t,s)|ds \le K_1 \quad \forall t > 0. \tag{4.1}$$

Proof. The proof follows as in [5]. We provide only the steps for the reader's convenience.

Define

$$\begin{aligned} (Sf)(t) &= \int_0^t X(t,\beta)f(\beta)d\beta, \quad f \in C_B[0,\infty), \\ S_k(f) &= \int_0^{t_k} X(t_k,\beta)f(\beta)d\beta, \quad f \in C_B[0,\infty), \end{aligned} \tag{4.2}$$

for each rational number t_k, $k \in \mathbb{N}$.

In view of (P), the family of continuous linear operators $\{S_k\}$ from $C_B[0,\infty)$ to $C_B[0,\infty)$ is pointwise-bounded. For the space of bounded continuous functions $C_B[0,\infty)$, the usual sup norm $\|\cdot\|$ is used.

By the Banach-Steinhaus theorem, the family is uniformly bounded. Thus, there is a positive number M such that $\|S_k(f)\| \le M\|f\|$ for every $f \in C_B[0,\infty)$.

As the rational numbers are dense in the real numbers, for each t there is t_k such that $t_k \to t$ as $k \to \infty$ and so

$$\left| \int_0^t X(t,\beta)f(\beta)d\beta \right| \le M\|f\| \quad \forall f \in C_B[0,\infty). \tag{4.3}$$

The final step involves choosing a sequence of functions and using a limiting process. □

Lemma 4.2. *If* (2.2) *and* (4.1) *are true, then there is a positive number K_2 such that*

$$|Y(s,t)| \le K_2 \quad \forall\, 0 \le s < t. \tag{4.4}$$

Proof. From (3.1), we have

$$Y(s,t) = I + \int_s^t A^T(\beta)Y(\beta,t)d\beta + \int_s^t B^T(h(\beta))Y(h(\beta),t)h'(\beta)d\beta. \tag{4.5}$$

to Contents, page 1

Hence, by using (4.1), we see that for all $0 \leq s < t$,

$$|Y(s,t)| \leq 1 + M_1 K_1 + M_2 K_1 =: K_2. \tag{4.6}$$

□

Lemma 4.3. *Let $G(r,t)$ be a continuous function satisfying $G(r,t) = 0$ for $r > t$. Then*

$$\int_{t_0}^{t} \left[\int_{s}^{h(s)} G(r,t)dr \right] ds = \int_{h(t_0)}^{t} (r - g(r))G(r,t)dr + \int_{t_0}^{h(t_0)} (r - t_0)G(r,t)dr. \tag{4.7}$$

5. Proofs of Theorems

Let $t_0 \geq 0$ be given. For a given continuous vector function ϕ defined on $[g(t_0), t_0]$, let $x(t) = x(t, t_0, \phi)$ denote the solution of (1.4) satisfying

$$x(t) = \phi(t), \quad t \leq t_0. \tag{5.1}$$

As usual,

$$\|\phi\|_g = \sup_{t \in [g(t_0), t_0]} |\phi(t)|. \tag{5.2}$$

Proof of Theorem 2.1. From Lemma 3.3, we may write

$$x(t) = Y^T(t_0, t)\phi(t_0) + \int_{g(t_0)}^{t_0} Y^T(h(\beta), t)B(h(\beta))\phi(\beta)h'(\beta)d\beta. \tag{5.3}$$

In view of Lemma 4.2, it follows that

$$|x(t)| \leq (K_2 + (h(t_0) - t_0)K_2 M_2)\|\phi\|_g. \tag{5.4}$$

Hence, the zero solution is stable. □

Proof of Theorem 2.2. Using (2.3) in (5.4), we get

$$|x(t)| \leq K_3\|\phi\|_g, \quad K_3 = K_2 + K_2 M_2 M_3, \tag{5.5}$$

from which the uniform stability follows. □

Proof of Theorem 2.3. By Theorem 2.1, the zero solution is stable. We need to show the attractivity property.

From Lemma 3.3, for $s \geq t_0$, we can write

$$x(t, t_0, \phi) = Y^T(s, t)x(s, t_0, \phi) + \int_{g(s)}^{s} G(h(\beta), t)x(\beta, t_0, \phi)h'(\beta)d\beta, \tag{5.6}$$

to Contents, page 1

where

$$G(s,t) = Y^T(s,t)B(s). \tag{5.7}$$

Integrating with respect to s from t_0 to t, we have

$$(t-t_0)x(t,t_0,\phi) = \int_{t_0}^{t}\left[Y^T(s,t)x(s,t_0,\phi) + \int_{s}^{h(s)} G(r,t)x(g(r),t_0,\phi)dr\right]ds. \tag{5.8}$$

We change the order of integration by employing Lemma 4.3. After some rearrangements, we obtain

$$\begin{aligned}(t-t_0)x(t,t_0,\phi) = {}&\int_{t_0}^{t} Y^T(s,t)x(s,t_0,\phi)ds + \int_{h(t_0)}^{t}(s-g(s))G(s,t)x(g(s),t_0,\phi)ds\\ &+\int_{t_0}^{h(t_0)}(s-t_0)G(s,t)x(g(s),t_0,\phi)ds.\end{aligned} \tag{5.9}$$

It follows that

$$(t-t_0)\left|x(t,t_0,\phi)\right| \le K_1K_3\|\phi\|_g + g_*(t,t_0)M_2K_1\|\phi\|_g + h_*(t_0)M_2K_1\|\phi\|_g. \tag{5.10}$$

In view of condition (2.4), we see from (5.10) that

$$\lim_{t\to\infty}\left|x(t,t_0,\phi)\right| = 0. \tag{5.11}$$

□

Proof of Theorem 2.4. By Theorem 2.2, the zero solution is uniformly stable. From (5.10) and (2.3), we have

$$(t-t_0)\left|x(t,t_0,\phi)\right| \le K_1K_3\|\phi\|_g + g_*(t,t_0)M_2K_1\|\phi\|_g + M_3M_2K_1\|\phi\|_g. \tag{5.12}$$

Using condition (2.4) in the above inequality, we see that the zero solution is uniformly asymptotically stable as $t \to \infty$. □

Acknowledgments

This research was supported by Grant P201/11/0768 of the Czech Grant Agency (Prague), by the Council of Czech Government MSM 0021630503 and MSM 00216 30519, and by Grant FEKT-S-11-2-921 of Faculty of Electrical Engineering and Communication, Brno University of Technology.

to Contents, page 1

References

[1] O. Perron, "Die stabilitätsfrage bei differentialgleichungen," *Mathematische Zeitschrift*, vol. 32, no. 1, pp. 703–728, 1930.

[2] R. Bellman, "On an application of a Banach-Steinhaus theorem to the study of the boundedness of solutions of non-linear differential and difference equations," *Annals of Mathematics*, vol. 49, pp. 515–522, 1948.

[3] J. Kloch, "An illustrative example for the Perron condition," *Annales Polonici Mathematici*, vol. 35, no. 1, pp. 11–14, 1978.

[4] R. Reissig, "A Perron-like stability criterion for linear systems," *Archiv der Mathematik*, vol. 34, no. 1, pp. 53–59, 1980.

[5] A. Halanay, *Differential Equations: Stability, Oscillations, Time Lags*, Academic Press, New York, NY, USA, 1966.

[6] M. U. Akhmet, J. Alzabut, and A. Zafer, "Perron's theorem for linear impulsive differential equations with distributed delay," *Journal of Computational and Applied Mathematics*, vol. 193, no. 1, pp. 204–218, 2006.

[7] A. Anokhin, L. Berezansky, and E. Braverman, "Exponential stability of linear delay impulsive differential equations," *Journal of Mathematical Analysis and Applications*, vol. 193, no. 3, pp. 923–941, 1995.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 354063, 17 pages
doi:10.1155/2011/354063

Research Article

On the Reducibility for a Class of Quasi-Periodic Hamiltonian Systems with Small Perturbation Parameter

Jia Li and Junxiang Xu

Department of Mathematics, Southeast University, Nanjing 210096, China

Correspondence should be addressed to Jia Li, lijia831112@163.com

Received 1 December 2010; Revised 18 April 2011; Accepted 25 May 2011

Academic Editor: Yuri V. Rogovchenko

We consider the following real two-dimensional nonlinear analytic quasi-periodic Hamiltonian system $\dot{x} = J\nabla_x H$, where $H(x,t,\varepsilon) = (1/2)\beta(x_1^2 + x_2^2) + F(x,t,\varepsilon)$ with $\beta \neq 0, \partial_x F(0,t,\varepsilon) = O(\varepsilon)$ and $\partial_{xx} F(0,t,\varepsilon) = O(\varepsilon)$ as $\varepsilon \to 0$. Without any nondegeneracy condition with respect to ε, we prove that for most of the sufficiently small ε, by a quasi-periodic symplectic transformation, it can be reduced to a quasi-periodic Hamiltonian system with an equilibrium.

1. Introduction

We first give some definitions and notations for our problem. A function $f(t)$ is called a quasi-periodic function with frequencies $\omega = (\omega_1, \omega_2, \ldots, \omega_l)$ if $f(t) = F(\omega_1 t, \omega_2 t, \ldots, \omega_l t)$ with $\theta_i = \omega_i t$, where $F(\theta_1, \theta_2, \ldots, \theta_l)$ is 2π periodic in all the arguments θ_j, $j = 1, 2, \ldots, l$. If $F(\theta)$ $(\theta = (\theta_1, \theta_2, \ldots, \theta_l))$ is analytic on $D_\rho = \{\theta \in C^l/2\pi Z^l \mid |\operatorname{Im}\theta_i| \leq \rho,\ i = 1, 2, \ldots, l\}$, we call $f(t)$ analytic quasi-periodic on D_ρ. If all $q_{ij}(t)$ $(i, j = 1, 2 \ldots, n)$ are analytic quasi-periodic on D_ρ, then the matrix function $Q(t) = (q_{ij}(t))_{1 \leq i, j \leq n}$ is called analytic quasi-periodic on D_ρ.

If $f(t)$ is analytic quasi-periodic on D_ρ, we can write it as Fourier series:

$$f(t) = \sum_{k \in Z^l} f_k e^{\mathrm{i}\langle k, \omega \rangle t}. \tag{1.1}$$

Define a norm of f by $\|f\|_\rho = \sum_{k \in Z^l} |f_k| e^{|k|\rho}$. It follows that $|f_k| \leq \|f\|_\rho e^{-|k|\rho}$. If the matrix function $Q(t)$ is analytic quasi-periodic on D_ρ, we define the norm of Q by $\|Q\|_\rho = n \times \max_{1 \leq i, j \leq n} \|q_{ij}\|_\rho$. It is easy to verify $\|Q_1 Q_2\|_\rho \leq \|Q_1\|_\rho \|Q_2\|_\rho$. The average of $Q(t)$ is denoted

to Contents, page 1

by $[Q] = ([q_{ij}])_{1\le i,j\le n}$, where

$$[q_{ij}] = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} q_{ij}(t)dt. \tag{1.2}$$

For the existence of the above limit, see [1].

Denote

$$D(r,\rho,\varepsilon_0) = \left\{(x,\theta,\varepsilon) \in C^n \times \left(\frac{C^l}{2\pi Z^l}\right) \times C \mid |x| \le r,\ \theta \in D_\rho,\ |\varepsilon| \le \varepsilon_0\right\}, \tag{1.3}$$

where $x = (x_1, x_2, \dots, x_n)$ and $|x| = |x_1| + |x_2| + \cdots + |x_n|$.

Let $f(x,t,\varepsilon)$ be analytic quasi-periodic of t and analytic in x and ε on $D(r,\rho,\varepsilon_0)$. Then $f(x,t,\varepsilon)$ can be expanded as

$$f(x,t,\varepsilon) = \sum_{m=0}^{\infty}\sum_{k\in Z^l} f_{mk}(x)\varepsilon^m e^{\mathrm{i}\langle k,\omega\rangle t}. \tag{1.4}$$

Define a norm by

$$\|f\|_{D(r,\rho,\varepsilon_0)} = \sum_{m=0}^{\infty}\sum_{k\in Z^l} |f_{mk}|_r \varepsilon_0^m e^{\rho|k|}, \tag{1.5}$$

where $|f_{mk}|_r = \sup_{|x|\le r}|f_{mk}(x)|$. Note that

$$\|f_1 \cdot f_2\|_{D(r,\rho,\varepsilon_0)} \le \|f_1\|_{D(r,\rho,\varepsilon_0)} \cdot \|f_2\|_{D(r,\rho,\varepsilon_0)}. \tag{1.6}$$

Problems

The reducibility on the linear differential system has been studied for a long time. The well-known Floquet theorem tells us that if $A(t)$ is a T-periodic matrix, then the linear system $\dot{x} = A(t)x$ is always reducible to the constant coefficient one by a T-periodic change of variables. However, this cannot be generalized to the quasi-periodic system. In [2], Johnson and Sell considered the quasi-periodic system $\dot{x} = A(t)x$, where $A(t)$ is a quasi-periodic matrix. Under some "full spectrum" conditions, they proved that $\dot{x} = A(t)x$ is reducible. That is, there exists a quasi-periodic nonsingular transformation $x = \phi(t)y$, where $\phi(t)$ and $\phi(t)^{-1}$ are quasi-periodic and bounded, such that $\dot{x} = A(t)x$ is transformed to $\dot{y} = By$, where B is a constant matrix.

In [3], Jorba and Simó considered the reducibility of the following linear system:

$$\dot{x} = (A + \varepsilon Q(t))x, \quad x \in R^n, \tag{1.7}$$

to Contents, page 1

where A is an $n \times n$ constant matrix with n different eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and $Q(t)$ is analytic quasi-periodic with respect to t with frequencies $\omega = (\omega_1, \omega_2, \ldots, \omega_l)$. Here ε is a small perturbation parameter. Suppose that the following nonresonance conditions hold:

$$\left| \langle k, \omega \rangle \sqrt{-1} + \lambda_i - \lambda_j \right| \geq \frac{\alpha}{|k|^\tau}, \tag{1.8}$$

for all $k \in Z^l \setminus \{0\}$, where $\alpha > 0$ is a small constant and $\tau > l - 1$. Assume that $\lambda_j^0(\varepsilon)$ ($j = 1, 2, \ldots, n$) are eigenvalues of $A + \varepsilon[Q]$. If the following non-degeneracy conditions hold:

$$\frac{d}{d\varepsilon} \left(\lambda_i^0(\varepsilon) - \lambda_j^0(\varepsilon) \right) \Big|_{\varepsilon=0} \neq 0, \quad \forall i \neq j, \tag{1.9}$$

then authors proved that for sufficiently small $\varepsilon_0 > 0$, there exists a nonempty Cantor subset $E \subset (0, \varepsilon_0)$, such that for $\varepsilon \in E$, the system (1.7) is reducible. Moreover, $\text{meas}((0, \varepsilon_0) \setminus E) = o(\varepsilon_0)$.

Some related problems were considered by Eliasson in [4, 5]. In the paper [4], to study one-dimensional linear Schrödinger equation

$$\frac{d^2 q}{dt^2} + Q(\omega t) q = Eq, \tag{1.10}$$

Eliasson considered the following equivalent two-dimensional quasi-periodic Hamiltonian system:

$$\dot{p} = (E - Q(\omega t)) q, \quad \dot{q} = p, \tag{1.11}$$

where Q is an analytic quasi-periodic function and E is an energy parameter. The result in [4] implies that for almost every sufficiently large E, the quasi-periodic system (1.11) is reducible. Later, in [5] the author considered the almost reducibility of linear quasi-periodic systems. Recently, the similar problem was considered by Her and You [6]. Let $C^\omega(\Lambda, gl(m, C))$ be the set of $m \times m$ matrices $A(\lambda)$ depending analytically on a parameter λ in a closed interval $\Lambda \subset R$. In [6], Her and You considered one-parameter families of quasi-periodic linear equations

$$\dot{x} = \left(A(\lambda) + g(\omega_1 t, \ldots, \omega_l t, \lambda) \right) x, \tag{1.12}$$

where $A \in C^\omega(\Lambda, gl(m, C))$, and g is analytic and sufficiently small. They proved that under some nonresonance conditions and some non-degeneracy conditions, there exists an open and dense set $\mathcal{A}$ in $C^\omega(\Lambda, gl(m, C))$, such that for each $A \in \mathcal{A}$, the system (1.12) is reducible for almost all $\lambda \in \Lambda$.

In 1996, Jorba and Simó extended the conclusion of the linear system to the nonlinear case. In [7], Jorba and Simó considered the quasi-periodic system

$$\dot{x} = (A + \varepsilon Q(t)) x + \varepsilon g(t) + h(x, t), \quad x \in R^n, \tag{1.13}$$

to Contents, page 1

where A has n different nonzero eigenvalues λ_i. They proved that under some nonresonance conditions and some non-degeneracy conditions, there exists a nonempty Cantor subset $E \subset (0, \varepsilon_0)$, such that the system (1.13) is reducible for $\varepsilon \in E$.

In [8], the authors found that the non-degeneracy condition is not necessary for the two-dimensional quasi-periodic system. They considered the two-dimensional nonlinear quasi-periodic system:

$$\dot{x} = Ax + f(x, t, \varepsilon), \quad x \in R^2, \tag{1.14}$$

where A has a pair of pure imaginary eigenvalues $\pm\sqrt{-1}\omega_0$ with $\omega_0 \neq 0$ satisfying the nonresonance conditions

$$|\langle k, \omega \rangle| \geq \frac{\alpha}{|k|^\tau}, \qquad |\langle k, \omega \rangle - 2\omega_0| \geq \frac{\alpha}{|k|^\tau} \tag{1.15}$$

for all $k \in Z^l \setminus \{0\}$, where $\alpha > 0$ is a small constant and $\tau > l - 1$. Assume that $f(0, t, \varepsilon) = O(\varepsilon)$ and $\partial_x f(0, t, \varepsilon) = O(\varepsilon)$ as $\varepsilon \to 0$. They proved that either of the following two results holds:

(1) for $\forall \varepsilon \in (0, \varepsilon_0)$, the system (1.14) is reducible to $\dot{y} = By + O(y)$ as $y \to 0$;

(2) there exists a nonempty Cantor subset $E \subset (0, \varepsilon_0)$, such that for $\varepsilon \in E$ the system (1.14) is reducible to $\dot{y} = By + O(y^2)$ as $y \to 0$.

Note that the result (1) happens when the eigenvalue of the perturbed matrix of A in KAM steps has nonzero real part. But the authors were interested in the equilibrium of the transformed system and obtained a small quasi-periodic solution for the original system.

Motivated by [8], in this paper we consider the Hamiltonian system and we have a better result.

2. Main Results

Theorem 2.1. *Consider the following real two-dimensional Hamiltonian system*

$$\dot{x} = J\nabla_x H, \quad x \in R^2, \tag{2.1}$$

where $H(x, t, \varepsilon) = (1/2)\beta(x_1^2 + x_2^2) + F(x, t, \varepsilon)$ *with* $\beta \neq 0$, $F(x, t, \varepsilon)$ *is analytic quasi-periodic with respect to* t *with frequencies* $\omega = (\omega_1, \omega_2, \ldots, \omega_l)$ *and real analytic with respect to* x *and* ε *on* $D(r, \rho, \varepsilon_0)$, *and*

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.2}$$

to Contents, page 1

Here $\varepsilon \in (0, \varepsilon_0)$ is a small parameter. Suppose that $\partial_x F(0,t,\varepsilon) = O(\varepsilon)$ and $\partial_{xx} F(0,t,\varepsilon) = O(\varepsilon)$ as $\varepsilon \to 0$. Moreover, assume that β and ω satisfy

$$|\langle k, \omega\rangle| \geq \frac{\alpha_0}{|k|^{\tau}}, \tag{2.3}$$

$$\left|\langle k, \omega\rangle - 2\beta\right| \geq \frac{\alpha_0}{|k|^{\tau}} \tag{2.4}$$

for all $k \in Z^l \setminus \{0\}$, where $\alpha_0 > 0$ is a small constant and $\tau > l - 1$.

Then there exist a sufficiently small $\varepsilon_* \in (0, \varepsilon_0]$ and a nonempty Cantor subset $E_* \subset (0, \varepsilon_*)$, such that for $\varepsilon \in E_*$, there exists an analytic quasi-periodic symplectic transformation $x = \phi_*(t)y + \psi_*(t)$ on $D_{\rho/2}$ with the frequencies ω, which changes (2.1) *into the Hamiltonian system $\dot{y} = J\nabla_y H_*$, where $H_*(y,t,\varepsilon) = 1/2\beta_*(\varepsilon)(y_1^2 + y_2^2) + F_*(y,t,\varepsilon)$, where $F_*(y,t,\varepsilon) = O(y^3)$ as $y \to 0$. Moreover,* $\mathrm{meas}((0,\varepsilon_*) \setminus E_*) = o(\varepsilon_*)$ *as $\varepsilon_* \to 0$. Furthermore, $\beta_*(\varepsilon) = \beta + O(\varepsilon)$ and $\|\phi_* - Id\|_{\rho/2} + \|\psi_*\|_{\rho/2} = O(\varepsilon)$, where Id is the 2-order unit matrix.*

3. The Lemmas

The proof of Theorem 2.1 is based on KAM-iteration. The idea is the same as [7, 8]. When the non-degeneracy conditions do not happen, the small parameter ε is not involved in the nonresonance conditions. So without deleting any parameter, the KAM step will be valid. Once the non-degeneracy conditions occur at some step, they will be kept for ever and we can apply the results with the non-degeneracy conditions. Thus, after infinite KAM steps, the transformed system is convergent to a desired form.

We first give some lemmas. Let $R = (r_{ij})_{1\leq i,j\leq 2}$ be a Hamiltonian matrix. Then we have $r_{11} + r_{22} = 0$. Define a matrix $R_A = (1/2)dJ$ with $d = r_{12} - r_{21}$. Let

$$B = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ \sqrt{-1} & -\sqrt{-1} \end{pmatrix}. \tag{3.1}$$

It is easy to verify

$$\begin{gathered} B^{-1}R_A B = \frac{1}{2}\operatorname{diag}\left(\sqrt{-1}d, -\sqrt{-1}d\right), \\ B^{-1}(R - R_A)B = \frac{1}{2}\begin{pmatrix} 0 & \sigma' - \sqrt{-1}\kappa' \\ \sigma' + \sqrt{-1}\kappa' & 0 \end{pmatrix}, \end{gathered} \tag{3.2}$$

where $\sigma' = 2r_{11}$ and $\kappa' = r_{21} + r_{12}$.

In the same way as in [7, 8], in KAM steps we need to solve linear homological equations. For this purpose we need the following lemma.

Lemma 3.1. *Consider the following equation of the matrix:*

$$\dot{P} = AP - PA + R(t), \tag{3.3}$$

to Contents, page 1

where $A = \beta(\varepsilon)J$ *with* $|\beta(\varepsilon)| > \mu$, $\mu > 0$ *is a constant, and* $R(t) = (r_{ij}(t))_{1\le i,j\le 2}$ *is a real analytic quasi-periodic Hamiltonian matrix on* D_ρ *with frequencies* ω. *Suppose* $\beta(\varepsilon)$ *and* R *are smooth with respect to* ε *and* $|\varepsilon\beta'(\varepsilon)| \le c_0$ *for* $\varepsilon \in E \subset (0, \varepsilon_*)$, *where* c_0 *is a constant. Note that here and below the dependence of* ε *is usually implied and one does not write it explicitly for simplicity. Assume* $[R]_A = 0$, *where* $[R]$ *is the average of* R. *Suppose that for* $\varepsilon \in E$, *the small divisors conditions* (2.3) *and the following small divisors conditions hold:*

$$\left|\langle k,\omega\rangle - 2\beta(\varepsilon)\right| \ge \frac{\alpha}{|k|^{\tau'}}, \tag{3.4}$$

where $\tau' > 2\tau + l$. *Let* $0 < s < \rho$ *and* $\rho_1 = \rho - s$. *Then there exists a unique real analytic quasi-periodic Hamiltonian matrix P(t) with frequencies* ω, *which solves the homological linear equation* (3.3) *and satisfies*

$$\|P\|_{\rho_1} \le \frac{c}{\alpha s^{v}}\|R\|_\rho, \qquad \|\varepsilon\partial_\varepsilon P\|_{\rho_1} \le \frac{c}{\alpha^2 s^{v'}}\left(\|R\|_\rho + \|\varepsilon\partial_\varepsilon R\|_\rho\right), \tag{3.5}$$

where $v = \tau' + l$, $v' = 2\tau' + l$ *and* $c > 0$ *is a constant.*

Remark 3.2. The subset E of $(0, \varepsilon_*)$ is usually a Cantor set and so the derivative with respect to ε should be understood in the sense of Whitney [9].

Proof. Let $\overline{P} = B^{-1}PB$, where B is defined by (3.1). Similarly, define $\overline{A}$, $\overline{R}$, $\overline{R}_A$. Then (3.3) becomes

$$\dot{\overline{P}} = \overline{A}\,\overline{P} - \overline{P}\,\overline{A} + \overline{R}(t), \tag{3.6}$$

where

$$\overline{A} = \operatorname{diag}\left(\sqrt{-1}\beta, -\sqrt{-1}\beta\right). \tag{3.7}$$

Moreover, $\overline{R}_A$ and $\overline{R} - \overline{R}_A$ have the same forms as (3.2) and (3.2), respectively

Noting that $[R]_A = 0$, we have $[\overline{R}]_A = 0$. Write $\overline{P} = (\overline{p}_{ij})_{i,j}$ and $\overline{R} = (\overline{r}_{ij})_{i,j}$. Obviously, we have $\overline{r}_{11} = -\overline{r}_{22}$ with $[\overline{r}_{ii}] = 0$.

Insert the Fourier series of $\overline{P}$ and $\overline{R}$ into (3.6). Then it follows that $\overline{p}_{ii}^0 = 0$, $\overline{p}_{ii}^k = \overline{r}_{ii}^k / (\langle k,\omega\rangle\sqrt{-1})$ for $k \ne 0$, and

$$\overline{p}_{ij}^k = \frac{\overline{r}_{ij}^k}{\sqrt{-1}\left(\langle k,\omega\rangle \pm 2\beta\right)} \quad \text{for } i \ne j. \tag{3.8}$$

Since $\overline{R}$ is analytic on D_ρ, we have $|\overline{R}_k| \le \|\overline{R}\|_\rho e^{-|k|\rho}$. So it follows

$$\left\|\overline{P}\right\|_{\rho-s} \le \sum_{k\in Z^l}\left|\overline{P}_k\right| e^{|k|\,(\rho-s)} \le \frac{c}{\alpha s^{v}}\|R\|_\rho. \tag{3.9}$$

to Contents, page 1

Note that here and below we always use c to indicate constants, which are independent of KAM steps.

Since A and $R(t)$ are real matrices, it is easy to obtain that $P(t)$ is also a real matrix. Obviously, it follows that $\overline{p}_{11} = -\overline{p}_{22}$ and the trace of the matrix $\overline{P}$ is zero. So is the trace of P. Thus, P is a Hamiltonian matrix.

Now we estimate $\|\varepsilon\partial P/\partial\varepsilon\|_{\rho_1}$. We only consider $\overline{p}_{12}$ and $\overline{p}_{21}$ since $\overline{p}_{11}$ and $\overline{p}_{22}$ are easy.

For $i \neq j$ we have

$$\frac{d\overline{p}_{ij}^{k}(\varepsilon)}{d\varepsilon} = \frac{\pm 2\beta'(\varepsilon)\overline{r}_{ij}^{k} - (\langle k,\omega\rangle \pm 2\beta)\overline{r}_{ij}^{k\,\prime}(\varepsilon)}{-\sqrt{-1}(\langle k,\omega\rangle \pm 2\beta)^2}. \tag{3.10}$$

Then, in the same way as above we obtain the estimate for $\|\varepsilon(\partial P/\partial\varepsilon)\|_{\rho_1}$. □

The following lemma will be used for the zero order term in KAM steps.

Lemma 3.3. *Consider the equation*

$$\dot{x} = Ax + g(t), \tag{3.11}$$

where A is the same as in Lemma 3.1, and g is real analytic quasi-periodic in t on D_ρ with frequencies ω and smooth with respect to ε. Suppose that the small divisors conditions (3.4) *hold. Then there exists a unique real analytic quasi-periodic solution $x(t)$ with frequencies ω, which satisfies*

$$\|x\|_{\rho_1} \le \frac{c}{\alpha s^{v}}\|g\|_\rho, \qquad \left\|\varepsilon\frac{\partial x}{\partial\varepsilon}\right\|_{\rho_1} \le \frac{c}{\alpha^2 s^{v'}}\left(\|g\|_\rho + \left\|\varepsilon\frac{\partial g}{\partial\varepsilon}\right\|_\rho\right), \tag{3.12}$$

where s,ρ_1,v, v' are defined in Lemma 3.1.

Proof. Similarly, let $\overline{x} = B^{-1}x$, $\overline{A} = B^{-1}AB$ and $\overline{g}(t) = B^{-1}g(t)$. Then (3.11) becomes

$$\dot{\overline{x}} = \overline{A}\,\overline{x} + \overline{g}(t), \tag{3.13}$$

where $\overline{A} = \operatorname{diag}(\sqrt{-1}\beta, -\sqrt{-1}\beta)$. Expanding $\overline{x} = (\overline{x}_1, \overline{x}_2)$ and $\overline{g} = (\overline{g}_1, \overline{g}_2)$ into Fourier series and using (3.13), we have

$$\overline{x}_i^{k} = \frac{\overline{g}_i^{k}}{\sqrt{-1}\left(\langle k,\omega\rangle + (-1)^i\beta\right)}. \tag{3.14}$$

to Contents, page 1

Using $2k$ in place of k in (3.4), we have

$$\left|\langle k,\omega\rangle - \beta(\varepsilon)\right| \geq \frac{\alpha}{2|k|^{\tau'}}. \tag{3.15}$$

Thus, in the same way as the proof of Lemma 3.1, we can estimate $\|x\|_{\rho_1}$ and $\|\varepsilon\partial_\varepsilon x\|_{\rho_1}$. We omit the details. □

The following lemma is used in the estimate of Lebesgue measure for the parameter ε in the case of non-degeneracy.

Lemma 3.4. *Let $\psi(\varepsilon) = \sigma\varepsilon^N + \varepsilon^N f(\varepsilon)$, where N is a positive integer and f satisfies that $f(\varepsilon) \to 0$ as $\varepsilon \to 0$ and $|f'(\varepsilon)| \leq c$ for $\varepsilon \in (0,\varepsilon_*)$. Let $\phi(\varepsilon) = \langle k,\omega\rangle - 2\beta - \psi(\varepsilon)$. Let*

$$O = \left\{\varepsilon \in (0,\varepsilon_*) \mid \left|\phi(\varepsilon)\right| \geq \frac{\alpha}{|k|^{\tau'}},\ \forall k \neq 0\right\}, \tag{3.16}$$

where $\tau' \geq 2\tau + l$, $\alpha \leq (1/2)\alpha_0$,$\sigma \neq 0$. Suppose that the small condition (2.4) *holds. Then when ε_* is sufficiently small, one has*

$$\operatorname{meas}(0,\varepsilon_*) \setminus O \leq c\frac{\alpha}{\alpha_0^2}\varepsilon_*^{N+1}, \tag{3.17}$$

where c is a constant independent of $\alpha_0, \alpha, \varepsilon_*$

Proof. Let

$$O_k = \left\{\varepsilon \in (0,\varepsilon_*) \mid \left|\phi(\varepsilon)\right| < \frac{\alpha}{|k|^{\tau'}}\right\}. \tag{3.18}$$

By assumption, if ε_* is sufficient small, we have that $|\psi(\varepsilon)| \leq 2\sigma\varepsilon^N$ and $|\psi'(\varepsilon)| \geq (\sigma/2)\varepsilon^{N-1}$ for $\varepsilon \in (0,\varepsilon_*)$. If $\varepsilon^N \leq \alpha_0/(4\sigma|k|^\tau)$, by (2.4) we have

$$\left|\phi(\varepsilon)\right| \geq \left|\langle k,\omega\rangle - 2\beta\right| - \left|\psi(\varepsilon)\right| \geq \frac{\alpha}{|k|^{\tau'}}. \tag{3.19}$$

Thus, we only consider the case that $\varepsilon_*^N \geq \varepsilon^N \geq (\alpha_0/(4\sigma|k|^\tau))$. We have $|k| \geq (\alpha_0/(4\sigma\varepsilon_*^N))^{1/\tau} = K$. Since

$$\left|\phi'(\varepsilon)\right| = \left|\psi'(\varepsilon)\right| \geq \frac{\sigma}{2}\varepsilon^{N-1} \geq \frac{\alpha_0}{8|k|^\tau\varepsilon_*}, \tag{3.20}$$

to Contents, page 1

we have $\text{meas}(O_k) \le ((2\alpha)/|k|^{\tau'}) \times ((8|k|^{\tau}\varepsilon_*)/\alpha_0) = (16\alpha\varepsilon_*)/(|k|^{\tau'-\tau}\alpha_0)$. So

$$\begin{aligned}\text{meas}((0,\varepsilon_*) \setminus 0) &\le \sum_{|k|\ge K} \text{meas}(O_k) \le \frac{16\alpha}{\alpha_0}\varepsilon_* \sum_{|k|\ge K} \frac{1}{|k|^{\tau'-\tau}} \\ &\le \frac{c\alpha}{\alpha_0}\varepsilon_* K^{l-\tau'+\tau} \le \frac{c\alpha}{\alpha_0^2}\varepsilon_*^{N+1},\end{aligned} \tag{3.21}$$

where c is a constant independent of α_0, α, and ε_*. □

Below we give a lemma with the non-degeneracy conditions.

Lemma 3.5. *Consider the real nonlinear Hamiltonian system $\dot{x} = J\nabla_x H$, where*

$$H(x,t,\varepsilon) = \frac{1}{2}\beta\left(x_1^2 + x_2^2\right) + F(x,t,\varepsilon) \quad \textit{with } \beta \ne 0. \tag{3.22}$$

Suppose that $F(x,t,\varepsilon)$ is analytic quasi-periodic with respect to t with frequencies ω and real analytic with respect to x and ε on $D(r,\rho,\varepsilon_0)$. Let $f(x,t,\varepsilon) = J\nabla_x F(x,t,\varepsilon)$. Assume that $f(0,t,\varepsilon) = O(\varepsilon^{2m_0})$ and $\partial_x f(0,t,\varepsilon) = O(\varepsilon^{m_0})$ as $\varepsilon \to 0$, where m_0 is a positive integer. Let $Q(t,\varepsilon) = \partial_x f(0,t,\varepsilon) = \sum_{k\ge m_0} Q_k(t)\varepsilon^k$. Suppose there exists $m_0 \le k \le 2m_0 - 1$ such that $[Q_k]_A \ne 0$ and the nonresonance conditions (2.3) *and* (2.4) *hold. Then, for sufficiently small $\varepsilon_* > 0$, there exists a nonempty Cantor subset $E_* \subset (0,\varepsilon_*)$, such that for $\varepsilon \in E_*$, there exists a quasi-periodic symplectic transformation $x = \phi_*(t)y + \psi_*(t)$ with the frequencies ω, which changes the Hamiltonian system to $\dot{y} = J\nabla_y H_*$, where*

$$H_*(y,t,\varepsilon) = \frac{1}{2}\beta_*(\varepsilon)\left(y_1^2 + y_2^2\right) + F_*(y,t,\varepsilon), \tag{3.23}$$

where $F_*(y,t,\varepsilon) = O(y^3)$ as $y \to 0$. Moreover, $\text{meas}((0,\varepsilon_*) \setminus E_*) = O(\varepsilon_*^{m_0+1})$ as $\varepsilon_* \to 0$. Furthermore, $\beta_*(\varepsilon) = \beta + O(\varepsilon^{m_0})$ and $\|\phi_* - Id\|_{\rho/2} + \|\psi_*\|_{\rho/2} = O(\varepsilon^{m_0})$.

Proof

KAM Step

The proof is based on a modified KAM iteration. In spirit, it is very similar to [7, 8]. The important thing is to make symplectic transformations so that the Hamiltonian structure can be preserved. Note that $[Q_k]_A \ne 0$ for some $m_0 \le k \le 2m_0 - 1$ is a non-degeneracy condition.

Consider the following Hamiltonian system

$$\dot{x} = Ax + f(x,t,\varepsilon), \tag{3.24}$$

where $A = \beta(\varepsilon)J$ and f is analytic quasi-periodic with respect to t with frequencies ω and real analytic with respect to x and ε on $D = D(r,\rho,\varepsilon_*)$.

to Contents, page 1

Let $\|f\|_D \le \alpha r\widetilde{\varepsilon}$ and $\|\varepsilon\partial_\varepsilon f\|_D \le \alpha r\widetilde{\varepsilon}$. Let $Q(t,\varepsilon) = \partial_x f(0,t,\varepsilon), g(t,\varepsilon) = f(0,t,\varepsilon)$ and

$$h(x,t,\varepsilon) = f(x,t,\varepsilon) - g(t,\varepsilon) - Q(t,\varepsilon)x. \tag{3.25}$$

Then h is the higher-order term of f. Moreover, the matrix $Q(t,\varepsilon)$ is Hamiltonian. Let $[Q]_A = \widehat{\beta}(\varepsilon)J$.

The system (3.24) is written as

$$\dot{x} = (A_+ + R(t,\varepsilon))x + g(t,\varepsilon) + h(x,t,\varepsilon), \tag{3.26}$$

where $A_+ = A + [Q]_A = \beta_+(\varepsilon)J$ and $R = Q - [Q]_A$. By assumption we have

$$\|g\|_\rho \le \alpha r\widetilde{\varepsilon}, \qquad \|Q\|_\rho \le \alpha\widetilde{\varepsilon}, \qquad \|h\|_D \le 3\alpha r\widetilde{\varepsilon}. \tag{3.27}$$

Moreover, we have

$$\|\varepsilon\partial_\varepsilon g\|_\rho \le \alpha r\widetilde{\varepsilon}, \qquad \|\varepsilon\partial_\varepsilon Q\|_\rho \le \alpha\widetilde{\varepsilon}, \qquad \|\varepsilon\partial_\varepsilon h\|_D \le 3\alpha r\widetilde{\varepsilon}. \tag{3.28}$$

Now we want to construct the symplectic change of variables $x = T'y = e^{P(t)}y$ to (3.26), where P is a Hamiltonian matrix to be defined later. Then we have

$$\begin{aligned}\dot{y} =& \left(e^{-P}(A_+ + R - \dot{P})e^P + e^{-P}\left(\dot{P}e^P - \frac{d}{dt}e^{P(t)}\right)\right)y\\ &+ e^{-P}g(t,\varepsilon) + e^{-P}h\left(e^P y,t,\varepsilon\right).\end{aligned} \tag{3.29}$$

Let $W = e^P - I - P$ and $\widetilde{W} = e^{-P} - I - P$. Then the system (3.29) becomes

$$\dot{y} = (A_+ + R - \dot{P} + A_+P - PA_+)y + Q'y + e^{-P}g(t,\varepsilon) + e^{-P}h\left(e^P y,t,\varepsilon\right), \tag{3.30}$$

where

$$\begin{aligned}Q' =& -P(R-\dot{P}) + (R-\dot{P})P - P(A_+ + R - \dot{P})P\\ &- P(A_+ + R - \dot{P})W + (A_+ + R - \dot{P})W\\ &+ \widetilde{W}(A_+ + R - \dot{P})e^P + e^{-P}\left(\dot{P}e^P - \frac{d}{dt}e^P\right).\end{aligned} \tag{3.31}$$

We would like to have

$$\dot{P} - A_+P + PA_+ = R, \tag{3.32}$$

to Contents, page 1

where $R = Q - [Q]_A$. Suppose the small divisors conditions (2.3) hold. Let $E_+ \subset (0, \varepsilon_*)$ be a subset such that for $\varepsilon \in E_+$ the small divisors conditions hold:

$$\left|\langle k, \omega\rangle - 2\beta_+(\varepsilon)\right| \geq \frac{\alpha_+}{|k|^{\tau'}}, \quad \forall k \in Z^l \setminus \{0\}, \tag{3.33}$$

where $\tau' > 2\tau + l$. By Lemma 3.1, we have a quasi-periodic Hamiltonian matrix $P(t)$ with frequencies ω to solve the above equation with the following estimates:

$$\begin{gathered} \|P\|_{\rho-s} \leq \frac{c\|Q\|_\rho}{\alpha_+ s^v} \leq \frac{c\tilde{\varepsilon}}{s^v}, \\ \left\|\varepsilon\frac{\partial P}{\partial \varepsilon}\right\|_{\rho-s} \leq \frac{c}{\alpha_+^2 s^{v'}}\left(\|Q\|_\rho + \left\|\varepsilon\frac{\partial Q}{\partial \varepsilon}\right\|_\rho\right) \leq \frac{c\tilde{\varepsilon}}{\alpha_+ s^{v'}}, \end{gathered} \tag{3.34}$$

where $v = \tau' + l, \ v' = 2\tau' + l$ and $c > 0$ is a constant. Then the system (3.30) becomes

$$\dot{y} = A_+ y + f'(y, t, \varepsilon), \tag{3.35}$$

where $f' = Q'y + e^{-P}g(t, \varepsilon) + e^{-P}h(e^P y, t, \varepsilon)$.

By Lemma 3.3, let us denote by $\underline{x}$ the solution of $\dot{x} = A_+ x + g'(t, \varepsilon)$ on $D_{\rho-2s}$, where $g' = e^{-P}g(t, \varepsilon)$. Then, by Lemma 3.3 we have

$$\begin{gathered} \|\underline{x}\|_{\rho-2s} \leq \frac{c\|g\|_{\rho-s}}{\alpha_+ s^v} \leq \frac{cr\tilde{\varepsilon}}{s^v}, \\ \left\|\varepsilon\frac{\partial \underline{x}}{\partial \varepsilon}\right\|_{\rho-2s} \leq \frac{c}{\alpha_+^2 s^{v'}}\left(\|g\|_{\rho-s} + \left\|\varepsilon\frac{\partial g}{\partial \varepsilon}\right\|_{\rho-s}\right) \leq \frac{cr\tilde{\varepsilon}}{\alpha_+ s^{v'}}. \end{gathered} \tag{3.36}$$

Under the symplectic change of variables $y = T'' x_+ = \underline{x} + x_+$, the Hamiltonian system (3.35) is changed to

$$\dot{x}_+ = A_+ x_+ + f_+(x_+, t, \varepsilon), \tag{3.37}$$

where $A_+ = \beta_+ J$ and

$$f_+ = Q' \cdot T'' + e^{-P} h \circ T' \circ T''. \tag{3.38}$$

to Contents, page 1

Let the symplectic transformation $T = T' \circ T''$. Then $x = Tx_+ = \phi(t)x_+ + \psi(t)$, where $\phi(t) = e^{P(t)}$ and $\psi(t) = e^{P(t)}\underline{x}(t)$. It is easy to obtain that if $\|P\|_{\rho-2s} \le 1/2$, then

$$\begin{aligned} \|\phi - I\|_{\rho-2s} &\le \frac{c\widetilde{\varepsilon}}{s^{v}}, \qquad \|\varepsilon\partial_\varepsilon\phi\|_{\rho-2s} \le \frac{c\widetilde{\varepsilon}}{\alpha_+ s^{v'}}, \\ \|\psi\|_{\rho-2s} &\le \frac{cr\widetilde{\varepsilon}}{s^{v}}, \qquad \|\varepsilon\partial_\varepsilon\psi\|_{\rho-2s} \le \frac{cr\widetilde{\varepsilon}}{\alpha_+ s^{v'}}. \end{aligned} \tag{3.39}$$

Under the symplectic change of variables $x = Tx_+$, the Hamiltonian system (3.24) becomes (3.37).

Below we give the estimates for A_+ and f_+. Obviously, it follows that $A_+(\varepsilon) - A = [Q]_A = \widehat{\beta}(\varepsilon)J$ and

$$\left|\beta_+(\varepsilon) - \beta(\varepsilon)\right| = \left|\widehat{\beta}(\varepsilon)\right| \le c\alpha\widetilde{\varepsilon}, \qquad \left|\varepsilon\left(\beta'_+(\varepsilon) - \beta'(\varepsilon)\right)\right| = \left|\varepsilon\widehat{\beta}'(\varepsilon)\right| \le c\alpha\widetilde{\varepsilon}. \tag{3.40}$$

By (3.38) we have

$$f_+(x_+, t, \varepsilon) = Q'(t)\left(x_+ + \underline{x}(t)\right) + e^{-P(t)}h\left(e^{P(t)}\left(x_+ + \underline{x}(t)\right), t, \varepsilon\right). \tag{3.41}$$

Let $\rho_+ = \rho - 2s$, and $r_+ = \eta r$ with $\eta \le 1/8$. If $c\widetilde{\varepsilon}/\alpha_+ s^{v+v'} \le \eta$, it follows that $\|\underline{x}\|_{\rho-2s} \le (1/8)r$. Let $D_+ = D(r_+, s_+, \varepsilon_*)$. Note that Q' and h only consist of high-order terms of P and x, respectively. It is easy to see $|e^{P(t)}(x_+ + \underline{x}(t))| \le 4\eta r \le r$. By all the estimates (3.27), (3.28), (3.34), and (3.36), and using usual technique of KAM estimate, we have

$$\begin{aligned} \|f_+\|_{D_+} &\le \frac{c\widetilde{\varepsilon}^2}{s^{2v}}\eta r + c\alpha r\widetilde{\varepsilon}\eta^2 \le \left(\frac{c\widetilde{\varepsilon}}{s^{2v}} + c\alpha\eta\right) r_+\widetilde{\varepsilon}, \\ \|\varepsilon\partial_\varepsilon f_+\|_{D_+} &\le \frac{c\widetilde{\varepsilon}^2}{\alpha_+ s^{v+v'}}\eta r + c\alpha r\widetilde{\varepsilon}\eta^2 \le \left(\frac{c\widetilde{\varepsilon}}{\alpha s^{v+v'}} + c\alpha\eta\right) r_+\widetilde{\varepsilon}. \end{aligned} \tag{3.42}$$

Let $\alpha_+ = \alpha/2$ and $\eta = c\widetilde{\varepsilon}/(\alpha^2 s^{v+v'})$. Then we have

$$\|f_+\|_{D_+} \le c\alpha_+ r_+ \eta\widetilde{\varepsilon} = \alpha_+ r_+ \widetilde{\varepsilon}_+, \qquad \widetilde{\varepsilon}_+ = c\eta\widetilde{\varepsilon}. \tag{3.43}$$

Similarly, we have

$$\|\varepsilon\partial_\varepsilon f_+\|_{D_+} \le \alpha_+ r_+ \widetilde{\varepsilon}_+. \tag{3.44}$$

Note that KAM steps only make sense for the small parameter ε satisfying small divisors conditions. However, by Whitney's extension theorem, for convenience all the functions are supposed to be defined for ε on $[0, \varepsilon_*]$.

to Contents, page 1

KAM Iteration

Now we can give the iteration procedure in the same way as in [7] and prove its convergence.

At the initial step, let $f_0 = f$. Let $f(x,t,\varepsilon) = f(0,t,\varepsilon) + \partial_x f(0,t,\varepsilon)x + h(x,t,\varepsilon)$. By assumption, if ε_* is sufficiently small, we have that for all $\varepsilon \in [0,\varepsilon_*]$

$$\begin{aligned} \left|f(0,t,\varepsilon)\right| &\le c\varepsilon^{2m_0}, & \left|\partial_x f(0,t,\varepsilon)\right| &\le c\varepsilon^{m_0}, \\ \left|\varepsilon\partial_\varepsilon f(0,t,\varepsilon)\right| &\le c\varepsilon^{2m_0}, & \left|\varepsilon\partial_\varepsilon\partial_x f(0,t,\varepsilon)\right| &\le c\varepsilon^{m_0}. \end{aligned} \tag{3.45}$$

Moreover,

$$|h(x,t,\varepsilon)| \le c|x|^2, \quad |\varepsilon\partial_\varepsilon h(x,t,\varepsilon)| \le c|x|^2, \quad \forall |x| \le \varepsilon^{m_0}, \quad \forall \varepsilon \in [0,\varepsilon_*]. \tag{3.46}$$

Let $r_0 = \varepsilon^{m_0}$, $\rho_0 = \rho$, $s_0 = \rho_0/8$, $D_0 = D(r_0,\rho_0,\varepsilon_*)$, and $\tilde{\varepsilon}_0 = c\varepsilon^{m_0}/\alpha_0$. Then we have

$$\left|f_0\right|_{D_0} \le \alpha_0 r_0 \tilde{\varepsilon}_0, \qquad \left|\varepsilon\partial_\varepsilon f_0\right|_{D_0} \le \alpha_0 r_0 \tilde{\varepsilon}_0. \tag{3.47}$$

For $n \ge 1$, let

$$\begin{aligned} \alpha_n &= \frac{\alpha_{n-1}}{2}, & s_n &= \frac{s_{n-1}}{2}, & \rho_n &= \rho_{n-1} - 2s_{n-1}, \\ \eta_{n-1} &= \frac{c\tilde{\varepsilon}_{n-1}}{\alpha_{n-1}^2 s_{n-1}^{v+v'}}, & r_n &= \eta_{n-1} r_{n-1}, & \tilde{\varepsilon}_n &= c\eta_{n-1}\tilde{\varepsilon}_{n-1}. \end{aligned} \tag{3.48}$$

Then we have a sequence of quasi-periodic symplectic transformations $\{T_n\}$ satisfying $T_n x = \phi_n(t)x + \varphi_n(t)$ with

$$\left\|\phi_n - I\right\|_{\rho_{n+1}} \le \frac{c\tilde{\varepsilon}_n}{s_n^v}, \qquad \left\|\varphi_n\right\|_{\rho_{n+1}} \le \frac{c r_n \tilde{\varepsilon}_n}{s_n^v}. \tag{3.49}$$

Let $T^n = T_0 \circ T_1 \cdots \circ T_{n-1}$. Then under the transformation $x = T^n y$ the Hamiltonian system $\dot{x} = A_0 x + f_0(x,t,\varepsilon)$ is changed to $\dot{y} = A_n y + f_n(y,t,\varepsilon)$.

Moreover, $A_n(\varepsilon) = \beta_n(\varepsilon)J$ satisfies $A_{n+1} - A_n = [Q_n]_A$ and

$$\left|\beta_{n+1}(\varepsilon) - \beta_n(\varepsilon)\right| \le c\alpha_n\tilde{\varepsilon}_n, \qquad \left|\varepsilon\left(\beta'_{n+1}(\varepsilon) - \beta'_n(\varepsilon)\right)\right| \le c\alpha_n\tilde{\varepsilon}_n, \tag{3.50}$$

$$\left\|f_n\right\|_{D_n} \le \alpha_n r_n \tilde{\varepsilon}_n. \tag{3.51}$$

Convergence

By the above definitions we have $\eta_n/\eta_{n-1} = c\tilde{\varepsilon}_n/\tilde{\varepsilon}_{n-1} = c\eta_{n-1}$. Thus, we have $\eta_n \le c\eta_{n-1}^2$ and so $c\eta_n \le (c\eta_{n-1})^2 \le (c\eta_0)^{2^n}$. Note that $\eta_0 = c\tilde{\varepsilon}_0/(\alpha_0^2 s_0^{v+v'}) \le c\varepsilon^{m_0}/(\alpha_0^2\rho_0^{v+v'})$. Suppose that ε_* is sufficiently small such that for $0 < \varepsilon < \varepsilon_*$ we have $c\eta_0 \le 1/2$. T_n are affine, so are T^n

to Contents, page 1

with $T^n x = \phi^n(t)x + \psi^n(t)$. By the estimates (3.49) it is easy to prove that $\phi^n(t)$ and $\psi^n(t)$ are convergent and so T^n is actually convergent on the domain $D(r/2, \rho/2)$. Let $T^n \to T_*$ and $T_* x = \phi_*(t)x + \psi_*(t)$. It is easy to see that the estimates for ϕ_* and ψ_* in Theorem 2.1 hold.

Using the estimate for f_n and Cauchy's estimate, we have $|f_n(0,t,\varepsilon)| \le \alpha_n r_n \widetilde{\varepsilon}_n \to 0$ and $|\partial_x f_n(0,t,\varepsilon)| \le \alpha_n \widetilde{\varepsilon}_n \to 0$ as $n \to \infty$. Let $f_n \to f_*$. Then it follows that $f_*(x,t,\varepsilon) = O(x^2)$.

By the estimates (3.50) for β_n we have $\beta_n \to \beta_*$. Thus, by the quasi-periodic symplectic transformation $x = T_* y$, the original system is changed to $\dot{y} = A_* y + f_*(y,t,\varepsilon)$ with $A_* = \beta_* J$.

Estimate of Measure

Let

$$E_n = \left\{ \varepsilon \in (0, \varepsilon_*) \mid \left| \langle \omega, k \rangle - 2\beta_n(\varepsilon) \right| \ge \frac{\alpha_n}{|k|^{\tau'}} \right\}. \tag{3.52}$$

Note that $\beta_n = \beta_1 + \psi$, where $\psi = \sum_{j=1}^{n-1} \beta_{j+1} - \beta_j, \beta_1 = \beta + \widehat{\beta}$, and $\widehat{\beta} J = [Q]_A$. Note that $\widetilde{\varepsilon}_1 = c\widetilde{\varepsilon}_0^2/(\alpha_0^2 s_0^{v+v'})$ and $\widetilde{\varepsilon}_0 = c\varepsilon^{m_0}/\alpha_0$. By the estimates (3.50), we have $\psi(\varepsilon) = O(\varepsilon^{2m_0})$ and $\varepsilon\psi'(\varepsilon) = O(\varepsilon^{2m_0})$. By assumption, $[Q]_A$ is analytic with respect to ε and there exists $m_0 \le N \le 2m_0 - 1$ such that $[Q]_A = \delta\varepsilon^N + O(\varepsilon^{N+1})$ with $\delta \ne 0$. Thus, $\beta_1(\varepsilon) = \beta + \delta\varepsilon^N + O(\varepsilon^{N+1})$. By Lemma 3.4, we have $\operatorname{meas}((0,\varepsilon_*) - E_n) \le c(\alpha_n/\alpha_0^2)\varepsilon_*^{N+1}$. Let $E_* = \bigcap_{n\ge 1} E_n$. By $\alpha_n = \alpha_0/2^n$, it follows that $\operatorname{meas}((0,\varepsilon_*) - E_*) \le c\varepsilon_*^{N+1}/\alpha_0$. Thus Lemma 3.5 is proved. □

4. Proof of Theorem 2.1

As we pointed previously, once the non-degeneracy conditions are satisfied in some KAM step, the proof is complete by Lemma 3.5. If the non-degeneracy conditions never happen, the small parameter ε does not involve into the small divisors and so the systems are analytic in ε. To prepare for KAM iteration, we need a preliminary step to change the original system to a suitable form.

Preliminary Step

We first give the preliminary KAM step. Let

$$\dot{x} = Ax + f(x,t,\varepsilon), \tag{4.1}$$

where $A = \beta J$ and $f = J\nabla_x F$. By Lemma 3.3, denote by $\underline{x}$ the solution of $\dot{x} = Ax + f(0,t,\varepsilon)$ on $D_{3\rho/4}$. Under the change of variables $x = T_0 x_+ = \underline{x} + x_+$, the Hamiltonian system (2.1) becomes

$$\dot{x}_+ = Ax_+ + f_1(x_+,t,\varepsilon), \tag{4.2}$$

where $f_1(x_+,t,\varepsilon) = f(\underline{x} + x_+,t,\varepsilon) - f(0,t,\varepsilon)$ satisfying $f_1(0,t,\varepsilon) = O(\varepsilon^2)$ and $\partial_{x_+} f_1(0,t,\varepsilon) = O(\varepsilon)$.

to Contents, page 1

KAM Step

The next step is almost the same as the proof of Lemma 3.5 and even more simple. In the KAM iteration, we only need consider the case that the non-degeneracy condition never happens. In this case, the normal frequency has no shift, which is equivalent to $A_n = A$ for all $n \geq 1$ in the KAM steps in the above nondegenerate case. Moreover, the small divisors conditions are always the initial ones as (2.3) and (2.4) and are independent of the small parameter ε. Thus, we need not delete any parameter. Moreover, the analyticity in ε remains in the KAM steps, which makes the estimate easier. At the first step, we consider $\dot{x} = Ax + f_1(x,t,\varepsilon)$. In the same way as the case of nondegenerate case, let $r_1 = \varepsilon, \rho_1 = 3\rho/4, \varepsilon_1 = \varepsilon_0, D_1 = D(r_1,\rho_1,\varepsilon_1)$, and $\widetilde{\varepsilon}_1 = c\varepsilon/\alpha_0$. Then we have $\|f_1\|_{D_1} \leq \alpha_0 r_1 \widetilde{\varepsilon}_1$.

At nth step, we consider the Hamiltonian system

$$\dot{x} = Ax + f_n(x,t,\varepsilon), \tag{4.3}$$

where f_n is analytic quasi-periodic with respect to t with frequencies ω and real analytic with respect to x and ε on $D_n = D(r_n,\rho_n,\varepsilon_n)$. Moreover, $\|f_n\|_{D_n} \leq \alpha_0 r_n \widetilde{\varepsilon}_n$. Suppose

$$Q_n(t,\varepsilon) = \partial_x f_n(0,t,\varepsilon) = O\left(\varepsilon^{2^{n-1}}\right), \quad f_n(0,t,\varepsilon) = O\left(\varepsilon^{2^n}\right). \tag{4.4}$$

Since Q_n is analytic with respect to ε, it follows that

$$Q_n = \sum_{k=2^{n-1}}^{\infty} Q_n^k \varepsilon^k. \tag{4.5}$$

Truncating the above power series of ε, we let

$$R_n(t,\varepsilon) = \sum_{k=2^{n-1}}^{2^n-1} Q_n^k \varepsilon^k, \qquad \widetilde{Q}_n = Q_n - R_n. \tag{4.6}$$

Because the non-degeneracy conditions do not happen in KAM steps, we must have $[R_n]_A = 0$. In the same way as the proof of Lemma 3.5, we have a quasi-periodic symplectic transformation T_n with $T_n x = \phi_n(t)x + \psi_n(t)$ satisfying (3.49). Let $T^n = T_1 \circ T_2 \cdots \circ T_{n-1}$.

By the transformation $x = T^n y$, the system (4.3) is changed to

$$\dot{y} = Ay + f_{n+1}(y,t,\varepsilon), \tag{4.7}$$

where $f_{n+1} = \widetilde{Q}_n \cdot T_n'' + Q_n' \cdot T_n'' + e^{-P_n} \cdot h_n \circ T_n = \widetilde{Q}_n(\underline{x}_n + y) + Q_n'(\underline{x}_n + y) + e^{-P_n} h_n(e^{P_n}(\underline{x}_n + y))$.

The last two terms can be estimated similarly as those of (3.41). Note that

$$\widetilde{Q}_n = Q_n - R_n = \sum_{k \geq 2^n} Q_n^k \varepsilon^k \tag{4.8}$$

to Contents, page 1

only consists of the higher order terms of ε. So, in the same way as [8, 10], we use the technique of shriek of the domain interval of ε to estimate the first term.

Let $r_1 = \varepsilon, \rho_1 = 3\rho/4, \varepsilon_1 = \varepsilon_0$ and $s_1 = \rho/16$.

Define $s_{n+1} = s_n/2, \rho_{n+1} = \rho_n - 2s_n, \eta_n = (1/8)e^{-(4/3)^n}$, $r_{n+1} = \eta_n r_n, \delta_n = 1 - (2/3)^n$ and $\varepsilon_{n+1} = \delta_n \varepsilon_n$. Let $D_{n+1} = D(r_{n+1}, \rho_{n+1}, \varepsilon_{n+1})$.

If $c\tilde{\varepsilon}_n/s_n^{2v} \leq \eta_n < (1/8)$, it follows that

$$\|f_{n+1}\|_{D_{n+1}} \leq \left(\alpha_0 \tilde{\varepsilon}_n e^{-(4/3)^n} + \left(\frac{c\tilde{\varepsilon}_n}{s_n^v}\right)^2\right)\eta_n r_n + c\alpha_0 r_n \tilde{\varepsilon}_n \eta_n^2 \leq \alpha_0 r_{n+1}\tilde{\varepsilon}_{n+1}, \tag{4.9}$$

where $\tilde{\varepsilon}_{n+1} = c\eta_n\tilde{\varepsilon}_n$. Moreover, it is easy to see

$$\partial_x f_{n+1}(0,t,\varepsilon) = O\left(\varepsilon^{2^n}\right), \quad f_{n+1}(0,t,\varepsilon) = O\left(\varepsilon^{2^{n+1}}\right). \tag{4.10}$$

Now we verify $c\tilde{\varepsilon}_n/s_n^{2v} \leq \eta_n < 1/8$. Let $G_n = c\tilde{\varepsilon}_n/s_n^{2v}$. By $G_n = ce^{-(4/3)^{n-1}}16^v G_{n-1}$, it follows that

$$G_n = (c16^v)^{n-1}e^{-[(4/3)^{n-1}+(4/3)^{n-2}+\cdots+(4/3)^1]}G_1 = (c16^v)^{n-1}e^4 e^{-4(4/3)^{n-1}}G_1. \tag{4.11}$$

Note that $G_1 = c\tilde{\varepsilon}_1/s_1^{2v}$. If $\tilde{\varepsilon}_1$ is sufficiently small, we have $c\tilde{\varepsilon}_n/s_n^{2v} = G_n \leq \eta_n$.

Note that $(cr_n\tilde{\varepsilon}_n/s_n^v) \to 0$ and $(c\tilde{\varepsilon}_n/(\eta_n s_n^v)) \to 0$ as $n \to \infty$, and $\tilde{\varepsilon}_n \leq cs_n^{2v}G_n$. Let $\varepsilon_* = \prod_{n\geq 1}(1-(2/3)^n)\varepsilon_0$. Thus, in the same way as before we can prove the convergence of the KAM iteration for all $\varepsilon \in (0, \varepsilon_*)$ and obtain the result of Theorem 2.1. We omit the details.

Remark 4.1. As suggested by the referee, we can also introduce an outer parameter to consider the Hamiltonian function $H(x,t,\varepsilon) = \langle\omega, I\rangle + (1/2)(\beta_* + \sigma(\varepsilon))(x_1^2 + x_2^2) + F(x,t,\varepsilon)$, where (θ, I) are the angle variable and the action variable and $x = (x_1, x_2)$ are a pair of normal variables. In the same way as in [11], $\sigma(\varepsilon)$ is the modified term of the normal frequency. Then by some technique as in [11–13], we can also prove Theorem 2.1.

Acknowledgments

The authors would like to thank the reviewers's suggestions about this revised version. This work was supported by the National Natural Science Foundation of China (11071038) and the Natural Science Foundation of Jiangsu Province (BK2010420).

References

[1] N. N. Bogoljubov, J. A. Mitropoliski, and A. M. Samoilenko, *Methods of Accelerated Convergence in Nonlinear Mechanics*, Springer, New York, NY, USA, 1976.

[2] R. A. Johnson and G. R. Sell, "Smoothness of spectral subbundles and reducibility of quasiperiodic linear differential systems," *Journal of Differential Equations*, vol. 41, no. 2, pp. 262–288, 1981.

to Contents, page 1

[3] A. Jorba and C. Simó, “On the reducibility of linear differential equations with quasiperiodic coefficients,” *Journal of Differential Equations*, vol. 98, no. 1, pp. 111–124, 1992.
[4] L. H. Eliasson, “Floquet solutions for the one-dimensional quasi-periodic Schrödinger equation,” *Communications in Mathematical Physics*, vol. 146, no. 3, pp. 447–482, 1992.
[5] L. H. Eliasson, “Almost reducibility of linear quasi-periodic systems,” in *Smooth Ergodic Theory and Its Applications (Seattle, WA, 1999)*, vol. 69 of *Proceedings of Symposia in Pure Mathematics*, pp. 679–705, American Mathematical Society, Providence, RI, USA, 2001.
[6] H.-L. Her and J. You, “Full measure reducibility for generic one-parameter family of quasi-periodic linear systems,” *Journal of Dynamics and Differential Equations*, vol. 20, no. 4, pp. 831–866, 2008.
[7] A. Jorba and C. Simó, “On quasi-periodic perturbations of elliptic equilibrium points,” *SIAM Journal on Mathematical Analysis*, vol. 27, no. 6, pp. 1704–1737, 1996.
[8] X. Wang and J. Xu, “On the reducibility of a class of nonlinear quasi-periodic system with small perturbation parameter near zero equilibrium point,” *Nonlinear Analysis*, vol. 69, no. 7, pp. 2318–2329, 2008.
[9] H. Whitney, “Analytic extensions of differentiable functions defined in closed sets,” *Transactions of the American Mathematical Society*, vol. 36, no. 1, pp. 63–89, 1934.
[10] L. Zhang and J. Xu, “Persistence of invariant torus in Hamiltonian systems with two-degree of freedom,” *Journal of Mathematical Analysis and Applications*, vol. 338, no. 2, pp. 793–802, 2008.
[11] J. Moser, “Convergent series expansions for quasi-periodic motions,” *Mathematische Annalen*, vol. 169, pp. 136–176, 1967.
[12] M. B. Sevryuk, “KAM-stable Hamiltonians,” *Journal of Dynamical and Control Systems*, vol. 1, no. 3, pp. 351–366, 1995.
[13] K. Soga, “A point-wise criterion for quasi-periodic motions in the KAM theory,” *Nonlinear Analysis*, vol. 73, no. 10, pp. 3151–3161, 2010.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 486714, 18 pages
doi:10.1155/2011/486714

Research Article

The Optimization of Solutions of the Dynamic Systems with Random Structure

Miroslava Růžičková[1] and Irada Dzhalladova[2]

[1] *University of Žilina, Žilina, Slovakia*
[2] *Vadim Getman Kyjew National Economic University, Kyjew, Ukraine*

Correspondence should be addressed to Miroslava Růžičková, miroslava.ruzickova@fpv.uniza.sk

Received 31 January 2011; Accepted 31 March 2011

Academic Editor: Josef Diblík

The paper deals with the class of jump control systems with semi-Markov coefficients. The control system is described as the system of linear differential equations. Every jump of the random process implies the random transformation of solutions of the considered system. Relations determining the optimal control to minimize the functional are derived using Lyapunov functions. Necessary conditions of optimization which enables the synthesis of the optimal control are established as well.

1. The Statement of the Problem

The optimal control theory as mathematical optimization method for deriving control policies plays an important role in the development of the modern mathematical control theory. The optimal control deals with the problem of finding such a control law for a given system that a certain optimality criterion is achieved. The background for the optimization method can be found in the work of Lev Pontryagin with his well-known Pontryagin's maximum principle. The optimal control has been applied in diverse fields, such as economics, bioengineering, process control, and many others. Some real-life problems are described by a continuous-time or discrete-time linear system of differential equations, but a lot of them are described by dynamic systems with random jumping changes, for example economics systems. The general theory of random structure systems can be found in the work of Artemiev and Kazakov [1]. The optimization of linear systems with random parameters are considered in many works, for example in [2–12]. Particularly, the original results concerning the stabilization of the systems with random coefficients and a random process are derived using moment equations and Lyapunov functions in [4]. These results create a more convenient technique for applying the method in practice using suitable software for engineering or economics investigation. Our aim is the expansion of the achieved results to a new

to Contents, page 1

class of systems of linear differential equations with semi-Markov coefficients and random transformation of solutions performed simultaneously with jumps of semi-Markov process. We will focus on using the particular values of Lyapunov functions for the calculation of coefficients of the control vector which minimize the quality criterion. We will also establish the necessary conditions of the optimal solution which enables the synthesis of the optimal control for the considered class of systems.

Let us consider the linear control system

$$\frac{dX(t)}{dt} = A(t,\xi(t))X(t) + B(t,\xi(t))U(t) \tag{1.1}$$

on the probability basis $(\Omega, \mathfrak{T}, \mathbf{P}, \mathbf{F} \equiv \{\mathbf{F}_t : t \geq 0\})$ and together with (1.1) we consider the initial conditions

$$X(0) = \varphi(\omega), \quad \varphi : \Omega \longrightarrow \mathbb{R}^n. \tag{1.2}$$

The coefficients of the system are semi-Markov coefficients defined by the transition intensities $q_{\alpha k}(t)$, $\alpha, k = 1, 2, \ldots, n$, from state θ_k to state θ_α. We suppose that the vectors $U(t)$ belong to the set of control U and the functions $q_{\alpha k}(t)$, $\alpha, k = 1, 2, \ldots, n$, satisfy the conditions [13]:

$$q_{\alpha k}(t) \geq 0, \quad \int_0^\infty q_k(t)dt = 1, \qquad q_k(t) \equiv \sum_{\alpha=1}^{\infty} q_{\alpha k}(t). \tag{1.3}$$

Definition 1.1. Let the matrices $Q(t,\xi(t))$, $L(t,\xi(t))$ with semi-Markov elements be symmetric and positive definite. The cost functional

$$J = \int_0^\infty \langle X^*(t)Q(t,\xi(t))X(t) + U^*(t)L(t,\xi(t))U(t)\rangle dt, \tag{1.4}$$

defined on the space $C^1 \times U$, where $\langle\cdot\rangle$ denotes mathematical expectation, is called *the quality criterion*.

Definition 1.2. Let $S(t,\xi(t))$ be a matrix with semi-Markov elements. The control vector

$$U(t) = S(t,\xi(t))X(t) \tag{1.5}$$

which minimizes the quality criterion $J(X,U)$ with respect to the system (1.1) is called *the optimal control*.

If we denote

$$\begin{gathered} G(t,\xi(t)) \equiv A(t,\xi(t)) + B(t,\xi(t))S(t,\xi(t)), \\ H(t,\zeta(t)) \equiv Q(t,\zeta(t)) + S^*(t,\zeta(t))L(t,\zeta(t))S(t,\zeta(t)), \end{gathered} \tag{1.6}$$

to Contents, page 1

then the system (1.1) can be rewritten to the form

$$\frac{dX(t)}{dt} = G(t, \xi(t))X(t), \tag{1.7}$$

and the functional (1.4) to the form

$$J = \int_0^{\infty} \langle X^*(t)H(t, \xi(t))X(t) \rangle dt. \tag{1.8}$$

We suppose also that, together with every jump of random process $\xi(t)$ in time t_j, the solutions of the system (1.7) submit to the random transformation

$$X(t_j + 0) = C_{sk}X(t_j - 0), \quad s, k = 1, 2, \ldots, n, \tag{1.9}$$

if the conditions $\xi(t_j + 0) = \theta_s$, $\xi(t_j - 0) = \theta_k$ hold.

Definition 1.3. Let $a_k(t)$, $k = 1, \ldots, n$, $t \geq 0$ be a selection of n different positive functions. If $\xi(t_j + 0) = \theta_s$, $\xi(t_j - 0) = \theta_k$, $s, k = 1, \ldots, n$, and for $t_j \leq t \leq t_{j+1}$ the equality $a(t, \xi(t) = \theta_s) = a_s(t - t_j)$ holds, then the function $a(t, \xi(t))$ is called *semi-Markov function*.

The application of semi-Markov functions makes it possible to use the concept of stochastic operator. In fact, the semi-Markov function $a(t, \xi(t))$ is an operator of the semi-Markov process $\xi(t)$, because the value of the semi-Markov function $a(t, \xi(t))$ is defined not only by the values t and $\xi(t)$, but it is also necessary to specify the function $a_s(t)$, $t \geq 0$ and the value of the jump of the process $\xi(t)$ in time t_j which precedes the moment of time t.

Our task is the construction of Lyapunov function for the new class of systems of linear differential equations with semi-Markov coefficients and then applying the function to solve the optimization problem which minimizes the quality criterion.

2. Auxiliary Results

In the proof of Theorem 3.1 in Section 3, we will employ two results concerning the construction of the Lyapunov function and the construction of the optimal control for the system of linear differential equations in a deterministic case. We will derive these auxiliary results in this part.

2.1. The Construction of the Lyapunov Function

Let us consider the system of linear differential equations

$$\frac{dX(t)}{dt} = A(t, \xi(t))X(t) \tag{2.1}$$

associated to the system (1.1).

to Contents, page 1

Let us define a quadratic form

$$w(t,x,\xi(t)) = x^* B(t,\xi(t))x, \qquad B(t,\xi(t)) > 0, \tag{2.2}$$

where elements of the matrix $B(t,\xi(t))$ are the semi-Markov processes. The matrix $B(t,\xi(t))$ is defined by such a set of n different symmetric and positive definite matrices $B_k(t)$, $t \geq 0$, $k = 1,\ldots,n$, that the equality $\xi(t) = \theta_s$ for $t_j \leq t \leq t_j + 1$ implies

$$B(t,\xi(t)) = B_s(t - t_j), \quad s = 1,2,\ldots,n. \tag{2.3}$$

Our purpose in this section is to express the value of the functional

$$v = \int_0^\infty \langle w(t,X(t),\xi(t))\rangle dt \tag{2.4}$$

in a convenient form, which can help us to prove the L_2-stability of the trivial solution of the system (2.1).

At first, we introduce the particular Lyapunov functions

$$v_k(x) = \int_0^\infty \langle w(t,X(t),\xi(t)) \mid X(t) = x,\ \xi(0) = \theta_k\rangle dt, \quad k = 1,2,\ldots,n. \tag{2.5}$$

If we can find the values of the particular Lyapunov functions in the form $v_k(x) = x^* C_k x$, $k = 1,2,\ldots,n$, then value of the functional v can be expressed by the formula

$$v = \int_{E_n} \sum_{k=1}^n v_k(x) f_k(0,x)dx = \sum_{k=1}^n \int_{E_n} C_k \circ xx^* f_k(0,x)dx = \sum_{k=1}^n C_k \circ D_k(0), \tag{2.6}$$

where the scalar value

$$N \circ S = \sum_{k=1}^l \sum_{j=1}^m v_{kj} s_{kj} \tag{2.7}$$

is called the scalar product of the two matrices $N = (v_{kj})$, $S = (s_{kj})$ and has the property [14]

$$\frac{D(N \circ S)}{DS} = N. \tag{2.8}$$

The first auxiliary result contains two equivalent, necessary, and sufficient conditions for the L_2-stability (see in [4]) of the trivial solution of the system (2.1) and one sufficient condition for the stability of the solutions.

to Contents, page 1

Theorem 2.1. *The trivial solution of the system* (2.1) *is* L_2*-stable if and only if any of the next two equivalent conditions hold:*

(1) *the system of equations*

$$C_k = H_k + \int_0^{\infty} \sum_{s=1}^{n} q_{sk}(t) N_k^*(t) C_{sk}^* C_s C_{sk} N_k(t) dt, \quad k = 1, 2, \ldots, n \tag{2.9}$$

has a solution $C_k > 0$, $k = 1, 2, \ldots, n$ *for* $H_k > 0$, $k = 1, 2, \ldots, n$,

(2) *the sequence of the approximations*

$$\begin{aligned} C_k^{(0)} &= 0, \\ C_k^{(j+1)} &= H_k + \int_0^{\infty} \sum_{s=1}^{n} q_{sk}(t) N_k^*(t) C_{sk}^* C_s^{(j)} C_{sk} N_k(t) dt, \quad k = 1, 2, \ldots, n, \ j = 0, 1, 2, \end{aligned} \tag{2.10}$$

converges.

Moreover, the solutions of the system (2.1) *are* L_2*-stabile, if there exist symmetric and positive definite matrices* $C_k > 0$, $k = 1, 2, \ldots, n$, *such that the property*

$$C_k - \int_0^{\infty} \sum_{s=1}^{n} q_{sk}(t) N_k^*(t) C_{sk}^* C_s C_{sk} N_k(t) dt > 0, \quad k = 1, 2, \ldots, n \tag{2.11}$$

holds.

Proof. We will construct a system of equations, which will define the particular Lyapunov functions $v_k(x)$, $k = 1, 2, \ldots, n$. Let us introduce the auxiliary semi-Markov functions

$$u_k(t, x) = \langle w(t, X(t), \xi(t)) \mid X(0) = x, \ \xi(0) = \theta_k \rangle, \quad k = 1, 2, \ldots, n. \tag{2.12}$$

For the state $\xi(t) = \theta_k$, $t \geq 0$ of the random process $\xi(t)$, the equalities

$$X(t) = N_k(t) x, \quad X(0) = x \tag{2.13}$$

are true. Simultaneously, with the jumps of the random process $\xi(t)$, the jumps of solutions of (2.1) occurred, so in view of (2.12), we derive the equations

$$u_k(t, x) = \psi_k(t) w_k(t, N_k(t) x) + \int_0^{t} \sum_{s=1}^{n} q_{sk}(\tau) u_s(t - \tau, C_{sk} N_k(\tau) x) d\tau, \quad k = 1, 2, \ldots, n. \tag{2.14}$$

Further, if we introduce denoting

$$u_k(t, x) = x^* u_k(t) x, \quad k = 1, 2, \ldots, n, \tag{2.15}$$

to Contents, page 1

then (2.14) can be rewritten as the system of integral equations for the matrix $u_k(t)$ in the form

$$u_k(t) = \varphi_k(t)N_k^*(t)B_k(t)N_k(t) + \int_0^t \sum_{s=1}^n q_{sk}(\tau)N_k^*(\tau)C_{sk}^* u_s(t-\tau)C_{sk}N_k(\tau)d\tau, \quad k = 1,2,\ldots,n. \tag{2.16}$$

We define matrices $C_k,\ k = 1,2,\ldots,n$ and functions $v_k(t),\ k = 1,2,\ldots,n$, with regard to (2.5) and (2.12), by formulas

$$C_k = \int_0^\infty u_k(t)dt, \qquad v_k(x) = \int_0^\infty u_k(t,x)dt. \tag{2.17}$$

Integrating the system (2.16) from 0 to ∞, we get the system

$$C_k = \int_0^\infty \varphi_k(t)N_k^*(t)B_k(t)N_k(t)dt + \int_0^\infty \sum_{s=1}^n q_{sk}(\tau)N_k^*(\tau)C_{sk}^* C_s C_{sk}N_k(\tau)d\tau, \quad k = 1,2,\ldots,n. \tag{2.18}$$

Similarly, integrating the system of (2.14), we get the system of equations determining the particular Lyapunov functions

$$v_k(x) = \int_0^\infty \varphi_k(t)w_k(t,N_k(t)x)dt + \int_0^\infty \sum_{s=1}^n q_{sk}(t)v_k(C_{sk}N_k(t)x)dt. \tag{2.19}$$

Let us denote

$$H_k = \int_0^\infty \varphi_k(t)N_k^*(t)B_k(t)N_k(t)dt, \quad k = 1,2,\ldots,n. \tag{2.20}$$

If there exist such positive constants λ_1, λ_2 that

$$\lambda_1 E \le B_k(t) \le \lambda_2 E, \tag{2.21}$$

or equivalent conditions

$$\lambda_1 \|x\|^2 \le x^* B_k(t)x \le \lambda_2 \|x\|^2 \tag{2.22}$$

hold, then the matrices H_k, $k = 1,2,\ldots,n$ are symmetric and positive definite. Using (2.17), the system (2.18) can be rewritten to the form

$$C_k = H_k + \sum_{s=1}^n L_{sk}^* C_s, \quad k = 1,2,\ldots,n. \tag{2.23}$$

to Contents, page 1

It is easy to see that the system (2.23) is conjugated to the system (2.9). Therefore, the existence of a positive definite solution $C_k > 0,\ k = 1, 2, \ldots, n$ of the system (2.23) is equivalent to the existence of a positive definite solution $B_k > 0,\ k = 1, 2, \ldots, n$ and it is equivalent to L_2-stability of the solution of the system (2.1). On the other hand, if the existence of the particular Lyapunov functions $v_k(x),\ k = 1, 2, \ldots, n$ in (2.5) implies L_2-stability of the solutions of the system (2.1), then, in view of conditions (2.22) and the convergence of the integral (2.17), we get the inequality

$$\int_0^\infty \langle w(t, X(t), \xi(t)) \rangle dt \geq \int_0^\infty \left\langle \|X\|^2 \right\rangle dt. \tag{2.24}$$

The theorem is proved. □

Remark 2.2. If the system of linear differential equations (2.1) is a system with piecewise constant coefficients and the function $w(t, X(t), \xi(t))$ has the form

$$w(t, X(t), \xi(t)) = x^* B(\xi(t)) x, \qquad B_k \equiv B(\theta_k), \quad k = 1, 2, \ldots, n, \tag{2.25}$$

then the system (2.18) can be written in the form

$$C_k = \int_0^\infty \psi_k(t) e^{A_k^* t} B_k e^{A_k t} dt + \int_0^\infty \sum_{s=1}^{n} q_{sk}(t) e^{A_k^* t} C_{sk}^* C_s C_{sk} e^{A_k t} dt, \quad k = 1, 2, \ldots, n. \tag{2.26}$$

Particularly, if the semi-Markov process $\xi(t)$ is identical with a Markov process, then the system (2.26) has the form

$$C_k = \int_0^\infty e^{a_{kk} t} e^{A_k^* t} B_k e^{A_k t} dt + \int_0^\infty \sum_{\substack{s=1 \\ s \neq k}}^{n} a_{sk} e^{a_{kk} t} e^{A_k^* t} C_{sk}^* C_s C_{sk} e^{A_k t} dt, \quad k = 1, 2, \ldots, n, \tag{2.27}$$

or, more simply

$$C_k = \int_0^\infty e^{a_{kk} t} e^{A_k^* t} \left(B_k + \sum_{\substack{s=1 \\ s \neq k}}^{n} a_{sk} C_{sk}^* C_s C_{sk} \right) e^{A_k t} dt, \quad k = 1, 2, \ldots, n. \tag{2.28}$$

Moreover, under the assumption that the integral in (2.28) converges, the system (2.28) is equivalent to the system of matrices equations

$$(E a_{kk} + A_k^*) C_k + C_k A_K + B_k + \sum_{\substack{s=1 \\ s \neq k}}^{n} a_{sk} C_{sk}^* C_s C_{sk} = 0, \quad k = 1, 2, \ldots, n, \tag{2.29}$$

to Contents, page 1

which can be written as the system

$$A_k^* C_k + C_k A_K + B_k + \sum_{\substack{s=1 \\ s\neq k}}^{n} a_{sk} C_{sk}^* C_s C_{sk} = 0, \quad k = 1, 2, \ldots, n, \tag{2.30}$$

if $C_{kk} = E$, $k = 1, 2, \ldots, n$.

Example 2.3. Let the semi-Markov process $\xi(t)$ take two states θ_1, θ_2 and let it be identical with the Markov process described by the system of differential equations

$$\begin{aligned} \frac{dp_1(t)}{dt} &= -\lambda p_1(t) + \lambda p_2(t), \\ \frac{dp_2(t)}{dt} &= \lambda p_1(t) - \lambda p_2(t). \end{aligned} \tag{2.31}$$

We will consider the L_2-stability of the solutions of the differential equation

$$\frac{dx(t)}{dt} = a(\xi(t))x(t), \qquad a(\theta_k) \equiv a_k, \tag{2.32}$$

constructing a system of the type (2.26) related to (2.32). The system is

$$c_1 = 1 + \int_0^\infty e^{2a_2 t} \lambda e^{-\lambda t} c_2 dt, \qquad c_2 = 1 + \int_0^\infty e^{2a_1 t} \lambda e^{-\lambda t} c_1 dt, \tag{2.33}$$

and its solution is

$$c_1 = \frac{(\lambda - a_1)(\lambda - 2a_2)}{2a_1 a_2 - \lambda(a_1 + a_2)}, \qquad c_2 = \frac{(\lambda - a_2)(\lambda - 2a_1)}{2a_1 a_2 - \lambda(a_1 + a_2)}. \tag{2.34}$$

The trivial solution of (2.32) is L_2-stable, if $c_1 > 0$ and $c_2 > 0$. Let the intensities of semi-Markov process $\xi(t)$ satisfy the conditions

$$q_{11}(t) \approx 0, \qquad q_{22}(t) \approx 0, \qquad q_{21}(t) - \lambda e^{-\lambda t} \approx 0, \qquad q_{12}(t) - \lambda e^{-\lambda t} \approx 0. \tag{2.35}$$

Then, using the Theorem 2.1, the conditions

$$\begin{aligned} &1 - c_1 \int_0^\infty q_{11}(t) e^{2a_1 t} dt - c_2 \int_0^\infty \left(q_{21}(t) - \lambda e^{-\lambda t}\right) e^{2a_2 t} dt > 0, \\ &1 - c_1 \int_0^\infty \left(q_{12}(t) - \lambda e^{-\lambda t}\right) e^{2a_1 t} dt - c_2 \int_0^\infty q_{22}(t) e^{2a_2 t} dt > 0 \end{aligned} \tag{2.36}$$

are sufficient conditions for the L_2-stability of solutions of (2.32).

to Contents, page 1

2.2. The Construction of an Optimal Control for the System of Linear Differential Equations in the Deterministic Case

Let us consider the deterministic system of the linear equations

$$\frac{dX(t)}{dt} = A(t)X(t) + B(t)U(t) \tag{2.37}$$

in the boundary field G, where $X \in \mathbb{R}^m$, $U \in \mathbb{R}^l$, and together with (2.37) we consider the initial conditions

$$X(t) = x_0. \tag{2.38}$$

We assume that the vector $U(t)$ belongs to the control set U. The quality criterion has the form of the quadratic functional

$$\begin{gathered} I(t) = \frac{1}{2}\int_t^{\infty} [X^*(\tau)C(\tau)X(\tau) + U^*(\tau)\mathbf{D}(\tau)C(\tau)]d\tau, \\ C^*(t) = C(t), \qquad \mathbf{D}^*(t) = \mathbf{D}(t) \end{gathered} \tag{2.39}$$

in the space $\mathbb{C}^1(G) \times U$. The control vector

$$U(t) = S(t)X(t), \quad \dim S(t) = l \times m, \tag{2.40}$$

which minimizes the quality criterion (2.39) is called the optimal control.

The optimization problem is the problem of finding the optimal control (2.40) from all feasible control U, or, in fact, it is the problem of finding the equation to determine $S(t)$, $\dim S(t) = l \times m$.

Theorem 2.4. *Let there exist the optimal control* (2.40) *for the system of* (2.37). *Then the control equations*

$$S = -\mathbf{D}^{-1}(t)B^*(t)\mathbf{\Psi}^*, \quad \mathbf{\Psi}^* = K(t)X(t), \tag{2.41}$$

where the matrix $K(t)$ satisfies the Riccati equation

$$\frac{dK(t)}{dt} = -C(t) - K(t)A(t) - A^*(t)K(t) + K^*(t)B(t)\mathbf{D}^{-1}(t)B^*(t)K(t), \tag{2.42}$$

determines the synthesis of the optimal control.

Proof. Let the control for the system (2.37) have the form (2.40), where the matrix $S(t)$ is unknown. Then, the minimum value of the quality criterion (2.39) is

$$\min_{S(t)} I(t) = \frac{1}{2}X^*(t)K(t)X(t) \equiv \nu(t, X(t)). \tag{2.43}$$

to Contents, page 1

Under assumption that the vector $X(t)$ is known and using Pontryagin's maximum principle [1, 15], the minimum of the quality criterion (2.39) is written as

$$\min_{S(t)} I(t) = \frac{1}{2}\boldsymbol{\Psi}(t)X(t), \quad \tau \geq t, \tag{2.44}$$

where

$$\boldsymbol{\Psi}(t) = \frac{Dv(t,x)}{Dx} = X^* K(t) \tag{2.45}$$

is the row-vector. If we take Hamiltonian function [15] of the form

$$H(t,x,U,\boldsymbol{\Psi}) = \boldsymbol{\Psi}(A(t)x + B(t)U) + \frac{1}{2}(x^* Cx + U^* \mathbf{D}U), \quad U = Sx, \tag{2.46}$$

the necessary condition for optimality is

$$\frac{\partial H}{\partial s_{kj}} = 0, \quad k = 1,2,\ldots,l, \; j = 1,2,\ldots,m, \tag{2.47}$$

where s_{kj} are elements of the matrix S. The scalar value

$$\frac{dH}{dS} = \left\| \frac{\partial H}{\partial s_{kj}} \right\|, \quad k = 1,2,\ldots,l, \; j = 1,2,\ldots,m, \tag{2.48}$$

is called derivative of the matrix H with respect to the matrix S.

Employing the scalar product of the two matrices in our calculation, the Hamiltonian function (2.46) can be rewritten into the form

$$H = \boldsymbol{\Psi} A(t)x + \frac{1}{2}x^* C(t)x + B^*(t)\boldsymbol{\Psi}^* x^* \circ S + \frac{1}{2}\mathbf{D}(t) \cdot Sxx^* \circ S, \tag{2.49}$$

and its derivative with respect to the matrix S is

$$\frac{dH}{dS} = B^*(t)\boldsymbol{\Psi}^* x^* + \mathbf{D}(t)Sxx^* = 0. \tag{2.50}$$

Because the equality (2.50) holds for any value of x, the expression of the vector control U has the form

$$U = Sx = -\mathbf{D}^{-1}(t)B^*(t)\boldsymbol{\Psi}^* = -\mathbf{D}^{-1}(t)B^*(t)K(t)x, \tag{2.51}$$

which implies

$$S = -\mathbf{D}^{-1}(t)B^*(t)\boldsymbol{\Psi}^*. \tag{2.52}$$

to Contents, page 1

If we put the expression of matrix S to (2.49), we obtain a new expression for the Hamiltonian function

$$H = \mathbf{\Psi}(t)A(t)x + \frac{1}{2}x^{*}C(t)x - \frac{1}{2}\mathbf{\Psi}B(t)\mathbf{D}^{-1}(t)B^{*}(t)\mathbf{\Psi}^{*}, \tag{2.53}$$

for which the canonical system of linear differential equations

$$\frac{dx}{dt} = \frac{DH}{D\mathbf{\Psi}}, \qquad \frac{d\mathbf{\Psi}}{dt} = \frac{DH}{Dx} \tag{2.54}$$

has the form

$$\begin{gathered} \frac{dx}{dt} = A(t)x - B(t)\mathbf{D}^{-1}(t)B^{*}(t)\mathbf{\Psi}^{*}, \\ \frac{d\mathbf{\Psi}^{*}}{dt} = -C(t)x - A^{*}(t)\mathbf{\Psi}^{*}. \end{gathered} \tag{2.55}$$

In the end, we define the matrix $K(t)$ as the integral manifolds of solutions of the system equations

$$\mathbf{\Psi}^{*} = K(t)X(t). \tag{2.56}$$

If we derive the system (2.56) with respect to t regarding the system (2.55) and extract the vector $\mathbf{\Psi}^{*}$, then we obtain the matrix differential equation (2.40). This equation is known as Riccati equation in literature, see for example in [16, 17]. The solution $K_T(t)$ of (2.42) satisfying the initial condition

$$K_T(t) = 0, \qquad T > 0 \tag{2.57}$$

determines the minimum of the functional

$$\min_{S(\tau)} \int_t^T [X^{*}(\tau)C(\tau)X(\tau) + U^{*}(\tau)\mathbf{D}(\tau)U(\tau)]d\tau = \frac{1}{2}X^{*}(t)K_T(t)X(t), \tag{2.58}$$

and $K(t)$ can be obtained as the limit of the sequence $\{K_T(t)\}_{T=1}^{\infty}$ of the successive approximations $K_T(t)$:

$$K(t) = \lim_{T\to\infty} K_T(t). \tag{2.59}$$

□

Remark 2.5. Similar results can be obtained from the Bellman equation [18], where the function $v(t,x)$ satisfie

$$\min_{S(t)}\left\{\frac{\partial v(t,x)}{\partial t} + \frac{Dv(t,x)}{Dx}[A(t) + B(t)S(t)]x + \frac{1}{2}x^{*}C(t)x + \frac{1}{2}x^{*}S^{*}(t)\mathbf{D}(t)S(t)x\right\} = 0. \tag{2.60}$$

to Contents, page 1

3. The Main Result

Theorem 3.1. *Let the coefficients of the control system* (1.1) *be the semi-Markov functions and let them be defined by the equations*

$$\frac{dX_k(t)}{dt} = G_k(t)X_k(t), \quad G_k(t) \equiv A_k(t) + B_k(t)S_k(t), \quad k = 1, \dots, n. \tag{3.1}$$

Then the set of the optimal control is a nonempty subset of the control U, which is identical with the family of the solutions of the system

$$U_s(t) = L_s^{-1}(t)B_s^*(t)R_s(t)X_s(t), \quad s = 1, \dots, n, \tag{3.2}$$

where the matrix $R_s(t)$ is defined by the system of Riccati type of differential equations

$$\begin{aligned}\frac{dR_s(t)}{dt} &= -Q_s(t) - A_s^*(t)R_s(t) - R_s(t)A_s(t) \\ &\quad + R_s(t)B_s(t)L_s^{-1}(t)B_s^*(t)R_s(t) - \frac{\Psi'_s}{\Psi_s(t)}R_s(t) \\ &\quad - \sum_{k=1}^{n}\frac{q_{ks}(t)}{\Psi_s(t)}C_{ks}^*R_k(0)C_{ks}, \quad s = 1, \dots, n.\end{aligned} \tag{3.3}$$

3.1. The Proof of Main Result Using Lyapunov Functions

It should be recalled that the coefficients of the systems (1.1), (1.7) and of the functionals (1.4), (1.8) have the form

$$\begin{gathered}A(t,\xi(t)) = A_s(t - t_j), \qquad B(t,\xi(t)) = B_s(t - t_j), \\ Q(t,\xi(t)) = Q_s(t - t_j), \qquad L(t,\xi(t)) = L_s(t - t_j), \qquad S(t,\xi(t)) = S_s(t - t_j),\end{gathered} \tag{3.4}$$

if $t_j \le t < t_{j+1}$, $\xi(t) = \theta_s$. In addition to this, we have

$$\begin{gathered}G(t,\xi(t)) = G_s(t - t_j) \equiv A_s(t - t_j) + B_s(t - t_j)S_s(t - t_j), \\ H(t,\xi(t)) = H_s(t - t_j) \equiv Q_s(t - t_j) + S_s^*(t - t_j)L_s(t - t_j)S_s(t - t_j).\end{gathered} \tag{3.5}$$

The formula

$$V = \sum_{k=1}^{n} C_k \circ D_k(0) = \sum_{k=1}^{n} \int_{E_m} \nu_k(x) f_k(0, x)\, dx \tag{3.6}$$

to Contents, page 1

is useful for the calculation of the particular Lyapunov functions $v_k(x) \equiv x^* C_k x$, $k = 1, \ldots, n$ of the functional (1.8). We get

$$v_k(x) \equiv x^* C_k x = \int_0^\infty \langle X^*(t) H(t, \xi(t)) X(t) \mid X(0) = x,\ \xi(0) = \theta_k \rangle dt, \quad k = 1, 2, \ldots, n, \tag{3.7}$$

or, the more convenient form

$$v_k(x) \equiv x^* C_k x = \int_0^\infty \left[X_k^*(t) \left(\Psi_k(t) Q_k(t) + \sum_{s=1}^{n} q_{sk}(t) C_{sk}^* C_s C_{sk} \right) U_k^*(t) \Psi_k(t) L_k(t) U_k(t) \right] dt, \quad k = 1, 2, \ldots, n. \tag{3.8}$$

Then the system (3.1) has the form

$$\frac{dX_k(t)}{dt} = A_k(t) X_k(t) + B_k(t) U_k(t), \quad U_k(t) \equiv S_k(t) X_k(t),\ k = 1, \ldots, n. \tag{3.9}$$

Let us assume that for the control system (1.1) the optimal control exists in the form (1.5) independent of the initial value $X(0)$. Regarding the formula (3.6), there exist minimal values of the particular Lyapunov functions $v_k(x)$, $k = 1, \ldots, n$, which are associated with the optimal control. It also follows from the fact that the functions $v_k(x)$, $k = 1, \ldots, n$ are particular values of the functional (3.6). Finding the minimal values $v_k(x)$, $k = 1, \ldots, n$ by choosing the optimal control $U_k(x)$ is a well-studied problem, for the main results see [16]. It is significant that all matrices C_s, $s = 1, \ldots, n$ of the integrand in the formula (3.8) are constant matrices, hence, solving the optimization problem they can be considered as matrices of parameters.

Therefore, the problem to find the optimal control (1.5) for the system (1.1) can be transformed to n problems to find the optimal control for the deterministic system (3.9), which is equivalent to the system of linear differential equations of type (2.37).

3.2. The Proof of the Main Result Using Lagrange Functions

In this part, we get one more proof of the Theorem 3.1 using the Lagrange function.

We are looking for the optimal control which reaches the minimum of quality criterion

$$x^* C x = \int_0^T [(X^*(t) Q A) X(t) + U^*(t) L(t) U(t)] dt. \tag{3.10}$$

Let us introduce the Lagrange function

$$I = \int_0^T \left[X^*(t) Q(t) X(t) + U^*(t) L(t) U(t) + 2Y^*(t) \left(A(t) X(t) + B(t) U(t) - \frac{dX(t)}{dt} \right) \right] dt, \tag{3.11}$$

to Contents, page 1

where $Y(t)$ is the column-vector of Lagrange multipliers. In accordance with Pontryagin's maximum principle, we put the first variations of the functionals ∂I_x, ∂I_y equal to zero and we obtain the system of linear differential equations

$$\begin{aligned} \frac{dX(t)}{dt} &= A(t)X(t) - B(t)L^{-1}(t)B^*(t)Y(t), \\ \frac{dY(t)}{dt} &= -Q(t)X(t) - A^*(t)Y(t). \end{aligned} \tag{3.12}$$

Then the optimal control $U(t)$ can be expressed by

$$U(t) = L^{-1}(t)B^*(t)Y(t), \quad Y(T) = 0. \tag{3.13}$$

The synthesis of the optimal control needs to find the integral manifolds of the solutions of the system (3.12) in the form

$$Y(t) = K(t)X(t), \quad K(T) = 0. \tag{3.14}$$

According to the theory of integral manifolds [19] we construct the differential matrix equations of the Riccati type

$$\frac{dK(t)}{dt} = -Q(t) - A^*(t)K(t)A(t) - K(t)B(t)L^{-1}(t)B^*(t)K(t). \tag{3.15}$$

for the matrix $K(t)$. Integrating them from time $t = T$ to time $t = 0$ and using the initial condition $K(T) = 0$ we obtain Lagrange functions for the optimal control

$$U(t) = -L^{-1}(t)B^*(t)K(t)X(t). \tag{3.16}$$

We will prove that

$$\int_t^T [X^*(\tau)Q(\tau)X(\tau) + U^*(\tau)L(\tau)U(\tau)]d\tau = X^*(t)K(t)X(t). \tag{3.17}$$

Differentiating the equality (3.17) with respect to t we obtain the matrix equation

$$\begin{aligned} -X^*(t)Q(t)X(t) - U^*(t)L(t)U(t) = X^*(t)\frac{dK(t)}{dt}X(t) + X^*(t)K(t)(A(t)X(t) + B(t)U(t)) \\ + (X^*(t)A^*(t) + U^*(t)B^*(t))K(t)X(t), \end{aligned} \tag{3.18}$$

and extracting the optimal control $U(t)$ we obtain differential equation for $K(t)$ identical with (3.15). The equality $K(t) = K^*(t)$ follows from the positive definite matrices $Q(t)$, $L(t)$ for $t < T$. Therefore, from (3.17) we get $K(t) = 0$; moreover, from (3.10) it follows that $C = K(0)$.

to Contents, page 1

Applying the formulas (3.15), (3.16) to the system (3.8) with minimal functionals (3.9), the expression for the optimal control can be found in the form

$$U_s(t) = -\Psi_s^{-1}(t)L_s^{-1}(t)B_s^*(t)K_s(t)X_s(t), \quad s = 1, 2, \ldots, n, \tag{3.19}$$

where symmetric matrices $K_s(t)$ satisfy the matrix system of differential equations

$$\begin{aligned}\frac{dK_s(t)}{dt} = {}& -\Psi_s(t) - Q_s(t) - A_s^*(t)K_s(t) - \sum_{k=1}^{n} q_{ks}C_{ks}^*C_kC_{ks} \\ & + K_s(t)B_s(t)\Psi_s^{-1}(t)L_s^{-1}(t)B_s(t)K_s(t) \quad s = 1, 2, \ldots, n.\end{aligned} \tag{3.20}$$

The systems (3.9), (3.20) define the necessary condition such that the solutions of the systems (1.4) will be optimal. In addition to this, the system (3.8) defines the matrices $S_k(t)$, $k = 1, 2, \ldots, n$, of the optimal control in the form

$$S_k(t) = -\Psi_k^{-1}(t)L_k^{-1}(t)B_k^*(t)K_k(t), \quad k = 1, 2, \ldots, n. \tag{3.21}$$

We define matrices C_s from the system equations (3.20) in the view of

$$C_s = K_s(0), \quad s = 1, 2, \ldots, n. \tag{3.22}$$

In regards to

$$R_s(t) = -\Psi_s^{-1}(t)K_s(t), \quad \Psi_s(0) = 1, \quad C_s = R_s(0), \quad s = 1, 2, \ldots, n, \tag{3.23}$$

it can makes the system (3.20) simpler. Then the system (3.20) takes the form (3.3), and formula (3.2) defines the optimal control.

Remark 3.2. If the control system (1.1) is deterministic, then $q_{ks}(t) \equiv 0$, $\Psi_s(t) \equiv 0$, $k, s = 1, 2, \ldots, n$ and the system (3.3) is identical to the system of the Riccati type equations (3.15).

4. Particular Cases

The optimal control $U(t)$ for the system (1.1) has some special properties, and the equations determining it are different from those given in the previous section in case the coefficients of the control system (1.1) have special properties or intensities $q_{sk}(t)$ satisfy some relations or some other special conditions are satisfied. Some of these cases will be formulated as corollaries.

Corollary 4.1. *Let the control system* (1.1) *with piecewise constant coefficients have the form*

$$\frac{dX(t)}{dt} = A(\xi(t))X(t) + B(\xi(t))U(t). \tag{4.1}$$

to Contents, page 1

Then the quadratics functional

$$V = \int_0^\infty \langle X^*(t)Q(\xi(t))X(t) + U^*(t)L(\xi(t))U(t)\rangle dt \tag{4.2}$$

determines the optimal control in the form

$$U(t) = S(t,\xi(t))X(t), \tag{4.3}$$

where

$$S(t,\xi(t)) = S_k\left(t - t_j\right), \tag{4.4}$$

and the matrices $S_k(t)$ satisfy the equations

$$S_k(t) = -L^{-1}B_k^* R_k(t), \quad k = 1,2,\ldots,n \tag{4.5}$$

if $t_j \le t < t_{j+1}$, $\xi(t) = \theta_k$.

The matrices $R_k(t)$, $k = 1,2,\ldots,n$ are the solutions of the systems of the Riccati-type equations:

$$\begin{aligned} \frac{dR_k(t)}{dt} &= -Q_k - A_k^* R_k(t) - R_k(t)A_k \\ &\quad + R_k(t)B_k L_k^{-1} B_k^* R_k(t) - \frac{\Psi'_k(t)}{\Psi_k(t)} R_k(t) \\ &\quad - \sum_{s=1}^n \frac{q_{sk}(t)}{\Psi_k(t)} C_{sk}^* R_s(0) C_{sk}, \quad k = 1,\ldots,n. \end{aligned} \tag{4.6}$$

Remark 4.2. In the corollary we mention piecewise constant coefficients of the control system (4.1). The coefficients of the functional (4.2) will be piecewise as well, but the optimal control is nonstationary.

Corollary 4.3. *Assume that*

$$\frac{\Psi'_k(t)}{\Psi_k(t)} = \text{const}, \quad \frac{q_{sk}(t)}{\Psi_k(t)} = \text{const}, \quad k,s = 1,2,\ldots,n. \tag{4.7}$$

Then the optimal control $U(t)$ will be piecewise constant.

Taking into consideration that the optimal control is piecewise constant, we find out that the matrices $R_k(t)$, $k = 1,2,\ldots,n$ in (4.5) are constant, which implies the form of the system (4.6) is changed to the form

$$Q_k + A_k^* R_k + R_k A_k - R_k B_k L_k^{-1} B_k^* R_k + \frac{\Psi'_k(t)}{\Psi_k(t)} R_k(t) + \sum_{s=1}^n \frac{q_{sk}(t)}{\Psi_k(t)} C_{sk}^* R_k C_{sk} = 0, \quad k = 1,\ldots,n. \tag{4.8}$$

to Contents, page 1

The system (4.8) has constant solutions R_k, $k = 1, 2, \ldots, n$, if conditions (4.7) hold. Moreover, if the random process $\xi(t)$ is a Markov process then the conditions (4.7) have the form

$$\frac{\Psi'_k(t)}{\Psi_k(t)} = a_{kk} = \text{const}, \quad \frac{q_{sk}(t)}{\Psi_k(t)} = a_{sk} = \text{const}, \quad k, s = 1, 2, \ldots, n, \ k \neq s, \tag{4.9}$$

and the system (4.8) transforms to the form

$$Q_k + A_k^* R_k + R_k A_k - R_k B_k L_k^{-1} B_k^* R_k + \sum_{s=1}^{n} a_{sk} C_{sk}^* R_s C_{sk} = 0, \quad k = 1, \ldots, n \tag{4.10}$$

for which the optimal control is

$$U(t) = S(\xi(t))X(t), \qquad S(\theta_k) \equiv S_k, \quad S_k = -L_k^{-1} B_k^* R_k, \quad k = 1, 2, \ldots, n. \tag{4.11}$$

Corollary 4.4. *Let the state θ_s of the semi-Markov process $\xi(t)$ be no longer than $T_s > 0$. Then the system* (3.8) *has the form*

$$\begin{aligned} v_k(x) &\equiv x^* C_k x \\ &= \int_0^{T_s} \left(X_k^*(t) \left(\Psi_k(t) Q_k(t) + \sum_{s=1}^{n} q_{sk}(t) C_{sk}^* C_s C_{sk} \right) X_k(t) + U_k^*(t) \Psi_k(t) L_k(t) U_k(t) \right) dt, \\ &\qquad k = 1, 2, \ldots, n. \end{aligned} \tag{4.12}$$

Because

$$K_s(T_s) = \Psi_s(t) R_s(t), \quad s = 1, 2, \ldots, n, \tag{4.13}$$

then

$$K_s(T_s) = 0, \quad s = 1, 2, \ldots, n. \tag{4.14}$$

In this case, the search for the matrix $K_s(t)$, $s = 1, 2, \ldots, n$ in concrete tasks is reduced to integration of the matrix system of differential equations (3.15) on the interval $[0, T_s]$ with initial conditions (4.14). In view of $\Psi_s(T_s) = 0$, $s = 1, 2, \ldots, n$, we can expect that every equation (3.15) has a singular point $t = T_s$. If $\Psi_s(t)$ has simple zero at the point $t = T_s$, then the system (4.6) meets the necessary condition

$$\Psi_s(T_s) R_s(T_s) + \sum_{k=1}^{n} q_{sk}(T_s) C_{ks}^* R_s(0) C_{ks} = 0, \quad s = 1, \ldots, n \tag{4.15}$$

for boundary of matrix $R_s(t)$ in the singular points.

to Contents, page 1

Acknowledgments

The authors would like to thank the following for their support: the Slovak Research and Development Agency (Project APPV-0700-07), the Grant Agency of the Slovak republic (VEGA1/0090/09), and the National Scholarship Program of the Slovak Republic (SAIA).

References

[1] V. M. Artemiev and I. E. Kazakov, *Handbook on the Theory of Automatic Control*, Nauka, Moscow, Russia, 1987.
[2] K.G. Valeev and I. A. Dzhalladova, *Optimization of Random Process*, KNEU, Kyjew, Russia, 2006.
[3] I. I. Gihman and A. V. Skorohod, *Controlable of Random Process*, Izdat. "Naukova Dumka", Kyjew, Russia, 1977.
[4] I. A. Dzhalladova, *Optimization of Stochastic System*, KNEU, Kyjew, Russia, 2005.
[5] V. K. Jasinskiy and E. V. Jasinskiy, *Problem of Stability and Stabilization of Dynamic Systems with Finite after Effect*, TVIMS, Kyjew, Russia, 2005.
[6] K. J. Astrom, *Introduction to Stochastic Control Theory*, vol. 70 of *Mathematics in Science and Engineering*, Academic Press, New York, NY, USA, 1970.
[7] R. Balaji, *Introduction to Stochastic Finance*, University to Conecticut, Academic Press, New York, NY, USA, 1997.
[8] J. K. Hale and S. M. Verduyn Lunel, *Introduction to Functional-Differential Equations*, vol. 99 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 1993.
[9] O. Hájek, *Control Theory in the Plane*, vol. 153 of *Lecture Notes in Control and Information Sciences*, Springer, Berlin, Germany, 1991.
[10] X. Liao and P. Yu, *Absolute Stability of Nonlinear Control Systems*, vol. 25 of *Mathematical Modelling: Theory and Applications*, Springer, New York, NY, USA, Second edition, 2008.
[11] G. D. Qushner, *Stochastic Stability and Control*, Mir, Moscow, Russia, 1969.
[12] L. Glass and M. C. Mackey, *From Clocks to Chaos. The Rythms of Life*, Princeton University Press, Princeton, NJ, USA, 1988.
[13] K.G. Valeev and O. L. Strijak, *Methods of Moment Equations*, AN USSR, Kyjew, Russia, 1985.
[14] K. G. Valeev, O. L. Karelova, and V. I. Gorelov, *Optimization of a System of Linear Differential Equations with Random Coefficients*, RUDN, Moscow, Russia, 1996.
[15] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, John Wiley & Sons, New York, NY, USA, 1972.
[16] K.G. Valeev and G. S. Finin, *Constructing of Lyapunov Function*, Naukova dumka, Kyjew, Russia, 1981.
[17] E. A. Barbashyn, *Lyapunov Functions*, Nauka, Moscow, Russia, 1970.
[18] R. Bellman, *Introduction to Matrix Analysis*, McGraw-Hill, New York, NY, USA, 1960.
[19] K. G. Valeev and O. A. Jautykov, *Infinite Systems of Differential Equations*, Nauka, Alma-Ata, Russia, 1974.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 387483, 9 pages
doi:10.1155/2011/387483

Research Article

Oscillation Criteria for Certain Second-Order Nonlinear Neutral Differential Equations of Mixed Type

Zhenlai Han,[1,2] Tongxing Li,[1,2] Chenghui Zhang,[2] and Ying Sun[1]

[1] *School of Science, University of Jinan, Jinan, Shandong 250022, China*
[2] *School of Control Science and Engineering, Shandong University, Jinan, Shandong 250061, China*

Correspondence should be addressed to Zhenlai Han, hanzhenlai@163.com

Received 19 September 2010; Accepted 19 January 2011

Academic Editor: Josef Diblík

Some oscillation criteria are established for the second-order nonlinear neutral differential equations of mixed type $[(x(t)+p_1x(t-\tau_1)+p_2x(t+\tau_2))^\gamma]'' = q_1(t)x^\gamma(t-\sigma_1)+q_2(t)x^\gamma(t+\sigma_2)$, $t \ge t_0$, where $\gamma \ge 1$ is a quotient of odd positive integers. Our results generalize the results given in the literature.

1. Introduction

This paper is concerned with the oscillatory behavior of the second-order nonlinear neutral differential equation of mixed type

$$\left[\left(x(t)+p_1x(t-\tau_1)+p_2x(t+\tau_2)\right)^\gamma\right]'' = q_1(t)x^\gamma(t-\sigma_1)+q_2(t)x^\gamma(t+\sigma_2), \quad t \ge t_0. \tag{1.1}$$

Throughout this paper, we will assume the following conditions hold.

(A_1) p_i, τ_i, and σ_i, $i=1,2$, are positive constants;

(A_2) $q_i \in C([t_0,\infty),[0,\infty))$, $i=1,2$.

By a solution of (1.1), we mean a function $x \in C([T_x,\infty),\mathbb{R})$ for some $T_x \ge t_0$ which has the property that $(x(t)+p_1x(t-\tau_1)+p_2x(t+\tau_2))^\gamma \in C^2([T_x,\infty),\mathbb{R})$ and satisfies (1.1) on $[T_x,\infty)$. As is customary, a solution of (1.1) is called oscillatory if it has arbitrarily large zeros on $[t_0,\infty)$, otherwise, it is called nonoscillatory. Equation (1.1) is said to be oscillatory if all its solutions are oscillatory.

to Contents, page 1

Neutral functional differential equations have numerous applications in electric networks. For instance, they are frequently used for the study of distributed networks containing lossless transmission lines which rise in high speed computers where the lossless transmission lines are used to interconnect switching circuits; see [1].

Recently, many results have been obtained on oscillation of nonneutral continuous and discrete equations and neutral functional differential equations, we refer the reader to the papers [2–35], and the references cited therein.

Philos [2] established some Philos-type oscillation criteria for the second-order linear differential equation

$$\left(r(t)x'(t)\right)' + q(t)x(t) = 0, \quad t \geq t_0. \tag{1.2}$$

In [3–5], the authors gave some sufficient conditions for oscillation of all solutions of second-order half-linear differential equation

$$\left(r(t)\left|x'(t)\right|^{\gamma-1}x'(t)\right)' + q(t)|x(\tau(t))|^{\gamma-1}x(\tau(t)) = 0, \quad t \geq t_0 \tag{1.3}$$

by employing a Riccati substitution technique.

Zhang et al. [15] examined the oscillation of even-order neutral differential equation

$$\left[x(t) + p(t)x(\tau(t))\right]^{(n)} + q(t)f(x(\sigma(t))) = 0, \quad t \geq t_0. \tag{1.4}$$

Some oscillation criteria for the following second-order quasilinear neutral differential equation

$$\left(r(t)\left|z'(t)\right|^{\gamma-1}z'(t)\right)' + q(t)|x(\sigma(t))|^{\gamma-1}x(\sigma(t)) = 0, \quad \text{for } z(t) = x(t) + p(t)x(\tau(t)),\ t \geq t_0 \tag{1.5}$$

were obtained by [12–17].

However, there are few results regarding the oscillatory properties of neutral differential equations with mixed arguments, see the papers [20–24]. In [25], the authors established some oscillation criteria for the following mixed neutral equation:

$$\left(x(t) + p_1x(t-\tau_1) + p_2x(t+\tau_2)\right)'' = q_1(t)x(t-\sigma_1) + q_2(t)x(t+\sigma_2), \quad t \geq t_0; \tag{1.6}$$

here q_1 and q_2 are nonnegative real-valued functions. Grace [26] obtained some oscillation theorems for the odd order neutral differential equation

$$\left(x(t) + p_1x(t-\tau_1) + p_2x(t+\tau_2)\right)^{(n)} = q_1x(t-\sigma_1) + q_2x(t+\sigma_2), \quad t \geq t_0, \tag{1.7}$$

to Contents, page 1

where $n \geq 1$ is odd. Grace [27] and Yan [28] obtained several sufficient conditions for the oscillation of solutions of higher-order neutral functional differential equation of the form

$$(x(t) + cx(t-h) + Cx(t+H))^{(n)} + qx(t-g) + Qx(t+G) = 0, \quad t \geq t_0, \tag{1.8}$$

where q and Q are nonnegative real constants.

Clearly, (1.6) is a special case of (1.1). The purpose of this paper is to study the oscillation behavior of (1.1).

In the sequel, when we write a functional inequality without specifying its domain of validity we assume that it holds for all sufficiently large t.

2. Main Results

In the following, we give our results.

Theorem 2.1. *Assume that $\sigma_i > \tau_i$, $i = 1, 2$. If*

$$\limsup_{t\to\infty} \int_t^{t+\sigma_2-\tau_2} (t+\sigma_2-\tau_2-s) Q_2(s)\,ds > \left(2^{\gamma-1}\right)^2 \left(1 + p_1^\gamma + \frac{p_2^\gamma}{2^{\gamma-1}}\right), \tag{2.1}$$

$$\limsup_{t\to\infty} \int_{t-\sigma_1+\tau_1}^{t} (s-t+\sigma_1-\tau_1) Q_1(s)\,ds > \left(2^{\gamma-1}\right)^2 \left(1 + p_1^\gamma + \frac{p_2^\gamma}{2^{\gamma-1}}\right), \tag{2.2}$$

where

$$Q_i(t) = \min\{q_i(t-\tau_1), q_i(t), q_i(t+\tau_2)\}, \tag{2.3}$$

for $i = 1, 2$, then every solution of (1.1) *oscillates.*

Proof. Let x be a nonoscillatory solution of (1.1). Without loss of generality, we assume that there exists $t_1 \geq t_0$ such that $x(t) > 0$, $x(t-\tau_1) > 0$, $x(t+\tau_2) > 0$, $x(t-\sigma_1) > 0$, and $x(t+\sigma_2) > 0$ for all $t \geq t_1$. Setting

$$\begin{aligned} z(t) &= \left(x(t) + p_1 x(t-\tau_1) + p_2 x(t+\tau_2)\right)^\gamma, \\ y(t) &= z(t) + p_1^\gamma z(t-\tau_1) + \frac{p_2^\gamma}{2^{\gamma-1}} z(t+\tau_2). \end{aligned} \tag{2.4}$$

Thus $z(t) > 0$, $y(t) > 0$, and

$$z''(t) = q_1(t) x^\gamma(t-\sigma_1) + q_2(t) x^\gamma(t+\sigma_2) \geq 0. \tag{2.5}$$

to Contents, page 1

Then, $z'(t)$ is of constant sign, eventually. On the other hand,

$$\begin{aligned} y''(t) &= q_1(t)x^\gamma(t-\sigma_1) + q_2(t)x^\gamma(t+\sigma_2) \\ &\quad + p_1^\gamma q_1(t-\tau_1)x^\gamma(t-\tau_1-\sigma_1) + p_1^\gamma q_2(t-\tau_1)x^\gamma(t-\tau_1+\sigma_2) \\ &\quad + \frac{p_2^\gamma}{2^{\gamma-1}} q_1(t+\tau_2)x^\gamma(t+\tau_2-\sigma_1) \\ &\quad + \frac{p_2^\gamma}{2^{\gamma-1}} q_2(t+\tau_2)x^\gamma(t+\tau_2+\sigma_2). \end{aligned} \tag{2.6}$$

Note that $g(u) = u^\gamma, \gamma \geq 1, u \in (0,\infty)$ is a convex function. Hence, by the definition of convex function, we obtain

$$a^\gamma + b^\gamma \geq \frac{1}{2^{\gamma-1}}(a+b)^\gamma. \tag{2.7}$$

Using inequality (2.7), we get

$$\begin{aligned} & x^\gamma(t-\sigma_1) + p_1^\gamma x^\gamma(t-\tau_1-\sigma_1) \geq \frac{1}{2^{\gamma-1}}\left(x(t-\sigma_1) + p_1 x(t-\tau_1-\sigma_1)\right)^\gamma, \\ & \frac{1}{2^{\gamma-1}}\left(x(t-\sigma_1) + p_1 x(t-\tau_1-\sigma_1)\right)^\gamma + \frac{p_2^\gamma}{2^{\gamma-1}} x^\gamma(t+\tau_2-\sigma_1) \\ & \quad \geq \frac{1}{\left(2^{\gamma-1}\right)^2}\left(x(t-\sigma_1) + p_1 x(t-\tau_1-\sigma_1) + p_2 x(t+\tau_2-\sigma_1)\right)^\gamma = \frac{z(t-\sigma_1)}{\left(2^{\gamma-1}\right)^2}. \end{aligned} \tag{2.8}$$

Similarly, we obtain

$$x^\gamma(t+\sigma_2) + p_1^\gamma x^\gamma(t-\tau_1+\sigma_2) + \frac{p_2^\gamma}{2^{\gamma-1}} x^\gamma(t+\tau_2+\sigma_2) \geq \frac{z(t+\sigma_2)}{\left(2^{\gamma-1}\right)^2}. \tag{2.9}$$

Thus, from (2.6), we have

$$y''(t) \geq \frac{1}{\left(2^{\gamma-1}\right)^2}\left(Q_1(t)z(t-\sigma_1) + Q_2(t)z(t+\sigma_2)\right). \tag{2.10}$$

In the following, we consider two cases.

Case 1. Assume that $z'(t) > 0$. Then, $y'(t) > 0$. In view of (2.10), we see that

$$y''(t+\tau_2) \geq \frac{1}{\left(2^{\gamma-1}\right)^2} Q_2(t+\tau_2) z(t+\tau_2+\sigma_2). \tag{2.11}$$

to Contents, page 1

Applying the monotonicity of z, we find

$$\begin{aligned} y(t+\sigma_2) &= z(t+\sigma_2) + p_1^\gamma z(t-\tau_1+\sigma_2) + \frac{p_2^\gamma}{2^{\gamma-1}} z(t+\tau_2+\sigma_2) \\ &\leq \left(1 + p_1^\gamma + \frac{p_2^\gamma}{2^{\gamma-1}}\right) z(t+\tau_2+\sigma_2). \end{aligned} \tag{2.12}$$

Combining the last two inequalities, we obtain the inequality

$$y''(t+\tau_2) \geq \frac{Q_2(t+\tau_2)}{\left(2^{\gamma-1}\right)^2 \left(1 + p_1^\gamma + p_2^\gamma / 2^{\gamma-1}\right)} y(t+\sigma_2). \tag{2.13}$$

Therefore, y is a positive increasing solution of the differential inequality

$$y''(t) \geq \frac{Q_2(t)}{\left(2^{\gamma-1}\right)^2 \left(1 + p_1^\gamma + p_2^\gamma / 2^{\gamma-1}\right)} y(t-\tau_2+\sigma_2). \tag{2.14}$$

However, by [11], condition (2.1) contradicts the existence of a positive increasing solution of inequality (2.14).

Case 2. Assume that $z'(t) < 0$. Then, $y'(t) < 0$. In view of (2.10), we see that

$$y''(t-\tau_1) \geq \frac{1}{\left(2^{\gamma-1}\right)^2} Q_1(t-\tau_1) z(t-\tau_1-\sigma_1). \tag{2.15}$$

Applying the monotonicity of z, we find

$$\begin{aligned} y(t-\sigma_1) &= z(t-\sigma_1) + p_1^\gamma z(t-\tau_1-\sigma_1) + p_2^\gamma \frac{1}{2^{\gamma-1}} z(t+\tau_2-\sigma_1) \\ &\leq \left(1 + p_1^\gamma + \frac{p_2^\gamma}{2^{\gamma-1}}\right) z(t-\tau_1-\sigma_1). \end{aligned} \tag{2.16}$$

Combining the last two inequalities, we obtain the inequality

$$y''(t-\tau_1) \geq \frac{Q_1(t-\tau_1)}{\left(2^{\gamma-1}\right)^2 \left(1 + p_1^\gamma + p_2^\gamma / 2^{\gamma-1}\right)} y(t-\sigma_1). \tag{2.17}$$

Therefore, y is a positive decreasing solution of the differential inequality

$$y''(t) \geq \frac{Q_1(t)}{\left(2^{\gamma-1}\right)^2 \left(1 + p_1^\gamma + p_2^\gamma / 2^{\gamma-1}\right)} y(t+\tau_1-\sigma_1). \tag{2.18}$$

to Contents, page 1

However, by [11], condition (2.2) contradicts the existence of a positive decreasing solution of inequality (2.18). □

Remark 2.2. When $\gamma = 1$, Theorem 2.1 involves results of [25, Theorem 1].

Theorem 2.3. *Let $\beta_i = (\sigma_i - \tau_i)/2 > 0$, $i = 1, 2$. Suppose that, for $i = 1, 2$, there exist functions*

$$a_i \in C^1[t_0, \infty), \qquad a_i(t) > 0, \qquad (-1)^i a_i'(t) \le 0, \tag{2.19}$$

such that

$$Q_i(t) \ge \left(2^{\gamma-1}\right)^2 \left(1 + p_1^\gamma + \frac{p_2^\gamma}{2^{\gamma-1}}\right) a_i(t) a_i\left(t + (-1)^i \beta_i\right), \tag{2.20}$$

where Q_i are as in (2.3) *for $i = 1, 2$. If the first-order differential inequality*

$$v'(t) + (-1)^{i+1} a_i\left(t + (-1)^i \beta_i\right) v\left(t + (-1)^i \beta_i\right) \ge 0 \tag{2.21}$$

has no eventually negative solution for $i = 1$ and no eventually positive solution for $i = 2$, then (1.1) *is oscillatory.*

Proof. Let x be a nonoscillatory solution of (1.1). Without loss of generality, we assume that there exists $t_1 \ge t_0$ such that $x(t) > 0$, $x(t-\tau_1) > 0$, $x(t+\tau_2) > 0$, $x(t-\sigma_1) > 0$, and $x(t+\sigma_2) > 0$ for all $t \ge t_1$. Define z and y as in Theorem 2.1. Proceeding as in the proof of Theorem 2.1, we get (2.10).

In the following, we consider two cases.

Case 1. Assume that $z'(t) > 0$. Clearly, $y'(t) > 0$. Then, just as in Case 1 of Theorem 2.1, we find that y is a positive increasing solution of inequality (2.14). Let $b_2(t) = y'(t) + a_2(t) y(t + \beta_2)$. Then $b_2(t) > 0$. Using (2.19) and (2.20), we obtain

$$\begin{aligned}
& b_2'(t) - \frac{a_2'(t)}{a_2(t)} b_2(t) - a_2(t) b_2(t + \beta_2) \\
&= y''(t) - \frac{a_2'(t)}{a_2(t)} y'(t) - a_2(t) a_2(t + \beta_2) y(t + 2\beta_2) \\
&\ge y''(t) - a_2(t) a_2(t + \beta_2) y(t + 2\beta_2) \\
&\ge y''(t) - \frac{Q_2(t)}{(2^{\gamma-1})^2 \left(1 + p_1^\gamma + \left(p_2^\gamma / 2^{\gamma-1}\right)\right)} y(t - \tau_2 + \sigma_2) \ge 0.
\end{aligned} \tag{2.22}$$

Define $b_2(t) = a_2(t) v(t)$. Then, v is a positive solution of (2.21) for $i = 2$, which is a contradiction.

to Contents, page 1

Case 2. Assume that $z'(t) < 0$. Clearly, $y'(t) < 0$. Then, just as in Case 2 of Theorem 2.1, we find that y is a positive decreasing solution of inequality (2.18). Let $b_1(t) = y'(t) - a_1(t)y(t-\beta_1)$. Then $b_1(t) < 0$. Using (2.19) and (2.20), we obtain

$$\begin{aligned} & b_1'(t) - \frac{a_1'(t)}{a_1(t)} b_1(t) + a_1(t) b_1(t-\beta_1) \\ & = y''(t) - \frac{a_1'(t)}{a_1(t)} y'(t) - a_1(t) a_1(t-\beta_1) y(t-2\beta_1) \\ & \geq y''(t) - a_1(t) a_1(t-\beta_1) y(t-2\beta_1) \\ & \geq y''(t) - \frac{Q_1(t)}{(2^{\gamma-1})^2 \left(1 + p_1^\gamma + p_2^\gamma / 2^{\gamma-1}\right)} y(t+\tau_1-\sigma_1) \geq 0. \end{aligned} \tag{2.23}$$

Define $b_1(t) = a_1(t)v(t)$. Then, v is a negative solution of (2.21) for $i = 1$. This contradiction completes the proof of the theorem. □

Remark 2.4. When $\gamma = 1$, Theorem 2.3 involves results of [25, Theorem 2].

From Theorem 2.3 and the results given in [12], we have the following oscillation criterion for (1.1).

Corollary 2.5. *Let* $\beta_i = (\sigma_i - \tau_i)/2 > 0$, $i = 1, 2$. *Assume that* (2.19) *and* (2.20) *hold for* $i = 1, 2$. *If*

$$\liminf_{t\to\infty} \int_{t-\beta_1}^{t} a_1(s-\beta_1) ds > \frac{1}{e}, \tag{2.24}$$

$$\liminf_{t\to\infty} \int_{t}^{t+\beta_2} a_2(s+\beta_2) ds > \frac{1}{e}, \tag{2.25}$$

then (1.1) *is oscillatory.*

Proof. It is known (see [12]) that condition (2.24) is sufficient for inequality (2.21) (for $i = 1$) to have no eventually negative solution. On the other hand, condition (2.25) is sufficient for inequality (2.21) (for $i = 2$) to have no eventually positive solution. □

For an application of our results, we give the following example.

Example 2.6. Consider the second-order differential equation

$$\left[\left(x(t) + p_1 x(t-\tau_1) + p_2 x(t+\tau_2)\right)^\gamma\right]'' = q_1 x^\gamma(t-\sigma_1) + q_2 x^\gamma(t+\sigma_2), \quad t \geq t_0, \tag{2.26}$$

where $q_i > 0$ are constants and $\sigma_i > \tau_i$ for $i = 1, 2$.

to Contents, page 1

It is easy to see that $Q_i(t) = q_i,\ i = 1, 2$. Assume that $\varepsilon > 0$. Let $a_i(t) = (2+\varepsilon)/(\mathrm{e}(\sigma_i - \tau_i)),\ i = 1, 2$. Clearly, (2.19) holds. If

$$q_i > \left[\frac{2}{(\mathrm{e}(\sigma_i - \tau_i))}\right]^2 \left(2^{\gamma-1}\right)^2 \left(1 + p_1^{\gamma} + \frac{p_2^{\gamma}}{2^{\gamma-1}}\right) \tag{2.27}$$

for $i = 1, 2$, then (2.20) holds. Moreover, we see that

$$\begin{aligned} \liminf_{t\to\infty} \int_{t-\beta_1}^{t} a_1(s-\beta_1)\mathrm{d}s &= \frac{2+\varepsilon}{2\mathrm{e}} > \frac{1}{\mathrm{e}}, \\ \liminf_{t\to\infty} \int_{t}^{t+\beta_2} a_2(s+\beta_2)\mathrm{d}s &= \frac{2+\varepsilon}{2\mathrm{e}} > \frac{1}{\mathrm{e}}. \end{aligned} \tag{2.28}$$

Hence by applying Corollary 2.5, we find that (2.26) is oscillatory.

Acknowledgments

The authors sincerely thank the referees for their constructive suggestions which improved the content of the paper. This research is supported by the Natural Science Foundation of China (11071143, 60904024, 11026112), China Postdoctoral Science Foundation funded project (200902564), and Shandong Provincial Natural Science Foundation (ZR2010AL002, ZR2009AL003, Y2008A28); it was also supported by University of Jinan Research Funds for Doctors (XBS0843).

References

[1] J. Hale, *Theory of Functional Differential Equations*, vol. 3 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 2nd edition, 1977.

[2] Ch. G. Philos, "Oscillation theorems for linear differential equations of second order," *Archiv der Mathematik*, vol. 53, no. 5, pp. 482–492, 1989.

[3] R. P. Agarwal, S.-L. Shieh, and C.-C. Yeh, "Oscillation criteria for second-order retarded differential equations," *Mathematical and Computer Modelling*, vol. 26, no. 4, pp. 1–11, 1997.

[4] J. Džurina and I. P. Stavroulakis, "Oscillation criteria for second-order delay differential equations," *Applied Mathematics and Computation*, vol. 140, no. 2-3, pp. 445–453, 2003.

[5] J. Baštinec, L. Berezansky, J. Diblík, and Z. Šmarda, "On the critical case in oscillation for differential equations with a single delay and with several delays," *Abstract and Applied Analysis*, vol. 2010, Article ID 417869, 20 pages, 2010.

[6] J. Baštinec, J. Diblík, and Z. Šmarda, "Oscillation of solutions of a linear second-order discrete-delayed equation," *Advances in Difference Equations*, vol. 2010, Article ID 693867, 12 pages, 2010.

[7] J. Diblík, Z. Svoboda, and Z. Šmarda, "Explicit criteria for the existence of positive solutions for a scalar differential equation with variable delay in the critical case," *Computers & Mathematics with Applications*, vol. 56, no. 2, pp. 556–564, 2008.

[8] L. Berezansky, J. Diblík, and Z. Šmarda, "Positive solutions of second-order delay differential equations with a damping term," *Computers & Mathematics with Applications*, vol. 60, no. 5, pp. 1332–1342, 2010.

[9] Y. G. Sun and F. W. Meng, "Note on the paper of J. Džurina and I. P. Stavroulakis," *Applied Mathematics and Computation*, vol. 174, no. 2, pp. 1634–1641, 2006.

[10] B. Baculíková, "Oscillation criteria for second order nonlinear differential equations," *Archivum Mathematicum*, vol. 42, no. 2, pp. 141–149, 2006.

to Contents, page 1

[11] R. G. Koplatadze and T. A. Chanturiya, *Ob ostsillyatsionnykh svoistvakh differentsialnykh uravnenii s otklonyayushchimsya argumentom (Oscillatory Properties of Differential Equations with Deviating Argument)*, Izdat. Tbilis. Univ., Tbilisi, Georgia, 1977.
[12] G. S. Ladde, V. Lakshmikantham, and B. G. Zhang, *Oscillation Theory of Differential Equations with Deviating Arguments*, vol. 110 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 1987.
[13] J. Diblík, Z. Svoboda, and Z. Šmarda, "Retract principle for neutral functional differential equations," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 12, pp. e1393–e1400, 2009.
[14] L. H. Erbe and Q. Kong, "Oscillation results for second order neutral differential equations," *Funkcialaj Ekvacioj*, vol. 35, no. 3, pp. 545–555, 1992.
[15] Q. Zhang, J. Yan, and L. Gao, "Oscillation behavior of even-order nonlinear neutral differential equations with variable coefficients," *Computers & Mathematics with Applications*, vol. 59, no. 1, pp. 426–430, 2010.
[16] Q. Wang, "Oscillation theorems for first-order nonlinear neutral functional differential equations," *Computers & Mathematics with Applications*, vol. 39, no. 5-6, pp. 19–28, 2000.
[17] Z. Han, T. Li, S. Sun, and Y. Sun, "Remarks on the paper [Appl. Math. Comput. 207 (2009) 388–396]," *Applied Mathematics and Computation*, vol. 215, no. 11, pp. 3998–4007, 2010.
[18] L. Liu and Y. Bai, "New oscillation criteria for second-order nonlinear neutral delay differential equations," *Journal of Computational and Applied Mathematics*, vol. 231, no. 2, pp. 657–663, 2009.
[19] R. Xu and F. Meng, "Oscillation criteria for second order quasi-linear neutral delay differential equations," *Applied Mathematics and Computation*, vol. 192, no. 1, pp. 216–222, 2007.
[20] J.-G. Dong, "Oscillation behavior of second order nonlinear neutral differential equations with deviating arguments," *Computers & Mathematics with Applications*, vol. 59, no. 12, pp. 3710–3717, 2010.
[21] J. Džurina and D. Hudáková, "Oscillation of second order neutral delay differential equations," *Mathematica Bohemica*, vol. 134, no. 1, pp. 31–38, 2009.
[22] M. Hasanbulli and Y. V. Rogovchenko, "Oscillation criteria for second order nonlinear neutral differential equations," *Applied Mathematics and Computation*, vol. 215, no. 12, pp. 4392–4399, 2010.
[23] B. Baculíková and J. Džurina, "Oscillation of third-order neutral differential equations," *Mathematical and Computer Modelling*, vol. 52, no. 1-2, pp. 215–226, 2010.
[24] S. H. Saker, "Oscillation of second order neutral delay differential equations of Emden-Fowler type," *Acta Mathematica Hungarica*, vol. 100, no. 1-2, pp. 37–62, 2003.
[25] J. Dzurina, J. Busha, and E. A. Airyan, "Oscillation criteria for second-order differential equations of neutral type with mixed arguments," *Differential Equations*, vol. 38, no. 1, pp. 137–140, 2002.
[26] S. R. Grace, "On the oscillations of mixed neutral equations," *Journal of Mathematical Analysis and Applications*, vol. 194, no. 2, pp. 377–388, 1995.
[27] S. R. Grace, "Oscillations of mixed neutral functional-differential equations," *Applied Mathematics and Computation*, vol. 68, no. 1, pp. 1–13, 1995.
[28] J. Yan, "Oscillations of higher order neutral differential equations of mixed type," *Israel Journal of Mathematics*, vol. 115, pp. 125–136, 2000.
[29] Z. Wang, "A necessary and sufficient condition for the oscillation of higher-order neutral equations," *The Tôhoku Mathematical Journal*, vol. 41, no. 4, pp. 575–588, 1989.
[30] Z. Han, T. Li, S. Sun, and W. Chen, "On the oscillation of second-order neutral delay differential equations," *Advances in Difference Equations*, vol. 2010, Article ID 289340, 8 pages, 2010.
[31] Z. Han, T. Li, S. Sun, C. Zhang, and B. Han, "Oscillation criteria for a class of second order neutral delay dynamic equations of Emden-Fowler type," *Abstract and Applied Analysis*, vol. 2011, Article ID 653689, 26 pages, 2011.
[32] T. Li, Z. Han, P. Zhao, and S. Sun, "Oscillation of even-order neutral delay differential equations," *Advances in Difference Equations*, vol. 2010, Article ID 184180, 9 pages, 2010.
[33] Z. Han, T. Li, S. Sun, and C. Zhang, "An oscillation criteria for third order neutral delay differential equations," *Journal of Applied Analysis*, vol. 16, no. 2, pp. 295–303, 2010.
[34] Z. Han, T. Li, S. Sun, and W. Chen, "Oscillation criteria for second-order nonlinear neutral delay differential equations," *Advances in Difference Equations*, vol. 2010, Article ID 763278, 23 pages, 2010.
[35] S. Sun, T. Li, Z. Han, and Y. Sun, "Oscillation of second-order neutral functional differential equations with mixed nonlinearities," *Abstract and Applied Analysis*, vol. 2011, Article ID 927690, 15 pages, 2011.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 653689, 26 pages
doi:10.1155/2011/653689

Research Article

Oscillation Criteria for a Class of Second-Order Neutral Delay Dynamic Equations of Emden-Fowler Type

Zhenlai Han,[1,2] Tongxing Li,[1,2] Shurong Sun,[1,3] Chao Zhang,[1] and Bangxian Han[4]

[1] *School of Science, University of Jinan, Jinan, Shandong 250022, China*
[2] *School of Control Science and Engineering, Shandong University, Jinan, Shandong 250061, China*
[3] *Department of Mathematics and Statistics, Missouri University of Science and Technology Rolla, Missouri 65409-0020, USA*
[4] *Department of Mathematics, University of Science and Technology, Hefei 230026, China*

Correspondence should be addressed to Zhenlai Han, hanzhenlai@163.com

Received 31 August 2010; Accepted 30 September 2010

Academic Editor: Elena Braverman

We establish some new oscillation criteria for the second-order neutral delay dynamic equations of Emden-Fowler type, $[a(t)(x(t)+r(t)x(\tau(t)))^{\Delta}]^{\Delta}+p(t)x^{\gamma}(\delta(t))=0$, on a time scale unbounded above. Here $\gamma>0$ is a quotient of odd positive integers with a and p being real-valued positive functions defined on $\mathbb{T}$. Our results in this paper not only extend and improve the results in the literature but also correct an error in one of the references.

1. Introduction

The study of dynamic equations on time scales, which goes back to its founder Hilger [1], is an area of mathematics that has recently received a lot of attention. It was partly created in order to unify the study of differential and difference equations. Many results concerning differential equations are carried over quite easily to corresponding results for difference equations, while other results seem to be completely different from their continuous counterparts. The study of dynamic equations on time scales reveals such discrepancies and helps avoid proving results twice—once for differential equations and once again for difference equations.

The three most popular examples of calculus on time scales are differential calculus, difference calculus, and quantum calculus (see Kac and Cheung [2]), that is, when

to Contents, page 1

$\mathbb{T} = \mathbb{R}$, $\mathbb{T} = \mathbb{N}$, and $\mathbb{T} = q^{\mathbb{N}_0} = \{q^t : t \in \mathbb{N}_0\}$, where $q > 1$. Many other interesting time scales exist, and they give rise to many applications (see [3]). Dynamic equations on a time scale have an enormous potential for applications such as in population dynamics. For example, it can model insect populations that are continuous while in season, die out in, for example, winter, while their eggs are incubating or dormant, and then hatch in a new season, giving rise to a nonoverlapping population (see [3]). There are applications of dynamic equations on time scales to quantum mechanics, electrical engineering, neural networks, heat transfer, and combinatorics. A recent cover story article in New Scientist [4] discusses several possible applications. Several authors have expounded on various aspects of this new theory; see the survey paper by Agarwal et al. [5] and references cited therein. A book on the subject of time scales, by Bohner and Peterson [3], summarizes and organizes much of time scale calculus; see also the book by Bohner and Peterson [6] for advances results of dynamic equations on time scales.

In recent years, there has been much research activity concerning the oscillation and nonoscillation of solutions of various dynamic equations on time scales unbounded above and neutral differential equations; we refer the reader to the papers [7–19]. Some authors are especially interested in obtaining sufficient conditions for the oscillation or nonoscillation of solutions of first and second-order linear and nonlinear neutral functional dynamic equations on time scales; we refer to the articles [20–28].

Agarwal et al. [7] considered the second-order delay dynamic equations

$$x^{\Delta\Delta}(t) + p(t)x(\tau(t)) = 0, \quad t \in \mathbb{T} \tag{1.1}$$

and established some sufficient conditions for oscillation of (1.1). Şahiner [11] studied the second-order nonlinear delay dynamic equations

$$x^{\Delta\Delta}(t) + p(t)f(x(\tau(t))) = 0, \quad t \in \mathbb{T} \tag{1.2}$$

and obtained some sufficient conditions for oscillation by employing Riccati transformation technique. Zhang and Zhu [13] examined the second-order dynamic equations

$$x^{\Delta\Delta}(t) + p(t)f(x(t-\tau)) = 0, \quad t \in \mathbb{T}, \tag{1.3}$$

and by using comparison theorems, they proved that oscillation of (1.3) is equivalent to the oscillation of the nonlinear dynamic equations

$$x^{\Delta\Delta}(t) + p(t)f(x(\sigma(t))) = 0, \quad t \in \mathbb{T} \tag{1.4}$$

and established some sufficient conditions for oscillation by applying the results established in [15]. Erbe et al. [16] investigated the oscillation of the second-order nonlinear delay dynamic equations

$$\left(r(t)x^{\Delta}(t)\right)^{\Delta} + p(t)f(x(\tau(t))) = 0, \quad t \in \mathbb{T} \tag{1.5}$$

to Contents, page 1

and by employing the generalized Riccati technique, they established some new sufficient conditions which ensure that every solution of (1.5) oscillates or converges to zero. Mathsen et al. [20] investigated the first-order neutral delay dynamic equations

$$[y(t) - r(t)y(\tau(t))]^{\Delta} + p(t)y(\delta(t)) = 0, \quad t \in \mathbb{T} \tag{1.6}$$

and established some new oscillation criteria which as a special case involve some well-known oscillation results for first-order neutral delay differential equations. Zhu and Wang [21] studied the nonoscillatory solutions to neutral dynamic equations

$$[y(t) + p(t)y(g(t))]^{\Delta} + f(t, x(h(t))) = 0, \quad t \in \mathbb{T} \tag{1.7}$$

and gave a classification scheme for the eventually positive solutions of (1.7). Agarwal et al. [22], Şahíner [23], Saker et al. [24–26], Wu et al. [27], and Zhang and Wang [28] considered the second-order nonlinear neutral delay dynamic equations

$$\left(r(t)\left((y(t) + p(t)y(\tau(t)))^{\Delta}\right)^{\gamma}\right)^{\Delta} + f(t, y(\delta(t))) = 0, \quad t \in \mathbb{T}, \tag{1.8}$$

where $\gamma > 0$ is a quotient of odd positive integers, the delay function τ and δ satisfy $\tau : \mathbb{T} \to \mathbb{T}$ and $\delta : \mathbb{T} \to \mathbb{T}$ for all $t \in \mathbb{T}$, and r and p are real-valued positive functions defined on $\mathbb{T}$, and

(h_1) $r(t) > 0, \int_{t_0}^{\infty} (1/r(t))^{1/\gamma} \Delta t = \infty$, and $0 \le p(t) < 1$;

(h_2) $f : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$ is continuous function such that $uf(u) > 0$ for all $u \neq 0$, and there exists a nonnegative function q defined on $\mathbb{T}$ such that $|f(t,u)| \ge q(t)|u|^{\gamma}$.

By employing different Riccati transformation technique, the authors established some oscillation criteria for all solutions of (1.8).

Recently, some authors have been interested in obtaining sufficient conditions for the oscillation and nonoscillation of solutions of Emden-Fowler type dynamic equations on time scales, differential equations, and difference equations; see, for example, [29–47].

Han et al. [32] studied the second-order Emden-Fowler delay dynamic equations

$$x^{\Delta\Delta}(t) + p(t)x^{\gamma}(\tau(t)) = 0, \quad t \in \mathbb{T} \tag{1.9}$$

and established some sufficient conditions for oscillation of (1.9) and extended the results given in [7].

Saker [34] studied the second-order superlinear neutral delay dynamic equation of Emden-Fowler type

$$\left[a(t)(y(t) + r(t)y(\tau(t)))^{\Delta}\right]^{\Delta} + p(t)\left|y(\delta(t))\right|^{\gamma} \operatorname{sign} y(\delta(t)) = 0 \tag{1.10}$$

on a time scale $\mathbb{T}$.

to Contents, page 1

The author assumes that

(A_1) $\gamma > 1$;

(A_2) the delay functions τ and δ satisfy $\tau : \mathbb{T} \to \mathbb{T}$, $\delta : \mathbb{T} \to \mathbb{T}$, $\tau(t) \le t, \delta(t) \le t$ for all $t \in \mathbb{T}$, and $\lim_{t\to\infty}\tau(t) = \lim_{t\to\infty}\delta(t) = \infty$;

(A_3) a, r and p are positive rd-continuous functions defined on $\mathbb{T}$ such that $a^{\Delta}(t) \ge 0$, $\int_{t_0}^{\infty}(\Delta t/a(t)) = \infty$, and $0 \le r(t) < 1$.

The main result for the oscillation of (1.10) in [34] is the following.

Theorem 1.1 (see, [34, Theorem 3.1]). *Assume that (A_1)–(A_3) hold. Furthermore, assume that*

$$\int_{t_0}^{\infty} p(t)(1 - r(\delta(t)))^{\gamma}\delta^{\gamma}(t)\Delta t = \infty, \tag{1.11}$$

and there exists a Δ-differentiable function η such that for all constants $M > 0$,

$$\limsup_{t\to\infty}\int_{t_0}^{t}\left[\eta(s)p(s)(1 - r(\delta(s)))^{\gamma}\left(\frac{\delta(s)}{s}\right)^{\gamma} - \frac{a(s)\left(\eta^{\Delta}(s)\right)^2}{4\gamma M^{\gamma-1}\eta(s)}\right]\Delta s = \infty. \tag{1.12}$$

Then every solution of (1.10) *is oscillatory.*

We note that in [34], the author gave an open problem, that is, how to establish oscillation criteria for (1.10) when $\gamma < 1$.

In [35], the author examined the oscillation of the second-order neutral delay dynamic equations

$$(x(t) - rx(\tau(t)))^{\Delta\Delta} + H(t, x(h_1(t))) = 0, \quad t \in \mathbb{T}. \tag{1.13}$$

The author assumes that

(H_1) τ and $h_1 \in C_{\mathrm{rd}}(\mathbb{T}, \mathbb{T}), \tau(t) < t$, $\tau(t) \to \infty$ as $t \to \infty$, $h_1(t) < t$, $h_1(t) \to \infty$ as $t \to \infty$, and $0 \le r < 1$;

(H_2) $H \in C(\mathbb{T} \times \mathbb{R}, \mathbb{R})$ for each $t \in \mathbb{T}$ which are nondecreasing in u, and $H(t, u) > 0$, for $u > 0$;

(H_3) $|H(t, u)| \ge \alpha(t)|u|^{\lambda}$, where $\alpha(t) \ge 0$, and $0 \le \lambda = p/q < 1$ with p, q being odd integers.

The main result for the oscillation of (1.13) in [35] is the following.

Theorem 1.2 (see, [35, Theorem 3.4]). *Assume that (H_1)–(H_3) hold. If for all sufficiently large $t_1 \ge t_0$,*

$$\int_{t_1}^{\infty}\alpha(s)(\tau(h_1(s)))^{\lambda}\Delta s = \infty, \tag{1.14}$$

then (1.13) *oscillates.*

to Contents, page 1

We find that the conclusion of this theorem is wrong. The following is a counter example of this theorem.

Counter Example. Consider the second-order differential equation

$$\left(x(t)-\frac{1}{3}x\left(\frac{t}{3}\right)\right)''+\left(\frac{1}{27}e^{-1/3}-e^{-1/3}e^{-2t/3}\right)x^{1/3}(t-1)=0,\quad t\geq t_0. \tag{1.15}$$

Let $\alpha(t)=e^{-1/3}/27-e^{-1/3}e^{-2t/3}$, $r(t)=1/3$, $\tau(t)=t/3$, and $h_1(t)=t-1$, $\lambda=1/3$. For all sufficiently large $t_1\geq t_0$, we find that

$$\int_{t_1}^{\infty}\alpha(s)(\tau(h_1(s)))^{\lambda}\Delta s=\int_{t_1}^{\infty}\alpha(s)(\tau(h_1(s)))^{\lambda}\mathrm{d}s=\int_{t_1}^{\infty}\left(\frac{1}{27}e^{-1/3}-e^{-1/3}e^{-2s/3}\right)\left(\frac{s-1}{3}\right)^{1/3}\mathrm{d}s. \tag{1.16}$$

It is easy to see that

$$\begin{gathered}\int_{t_1}^{\infty}\frac{1}{27}e^{-1/3}\left(\frac{s-1}{3}\right)^{1/3}\mathrm{d}s=\infty,\\ \int_{t_1}^{\infty}e^{-2s/3}\left(\frac{s-1}{3}\right)^{1/3}\mathrm{d}s\leq\int_{t_1}^{\infty}e^{-2s/3}s^{1/3}\mathrm{d}s.\end{gathered} \tag{1.17}$$

Integrating by parts, we obtain

$$\int_{t_1}^{\infty}e^{-2s/3}s^{1/3}\mathrm{d}s=-t_1^{1/3}\left(\frac{3}{2}e^{-2t_1/3}\right)+\frac{1}{2}\int_{t_1}^{\infty}e^{-2s/3}s^{-2/3}\mathrm{d}s<\infty. \tag{1.18}$$

Hence

$$\int_{t_1}^{\infty}\alpha(s)(\tau(h_1(s)))^{\lambda}\mathrm{d}s=\infty. \tag{1.19}$$

Therefore, by the above theorem, (1.15) is oscillatory. However, $x(t)=\mathrm{e}^{-t}$ is a positive solution of (1.15). Therefore, the above theorem is wrong. Tracing the error to its source, we find that the following false assertion was used in the proof of the aforementioned theorem.

Assertion A

If x is an eventually positive solution of (1.13) , then $z(t)=x(t)-r(t)x(\tau(t))$ is eventually positive.

Abdalla [37] studied the second-order superlinear neutral delay differential equations

$$\left[a(t)\left(y(t)+r(t)y(\tau(t))\right)'\right]'+p(t)\left|y(\delta(t))\right|^{\gamma}\operatorname{sign}y(\delta(t))=0,\quad t\in[t_0,\infty). \tag{1.20}$$

to Contents, page 1

Most of the oscillation criteria are unsatisfactory since additional assumptions have to be imposed on the unknown solutions. Also, the author proved that if

$$\int_{t_0}^{\infty} \frac{\mathrm{d}t}{a(t)} = \int_{t_0}^{\infty} p(t)\mathrm{d}t = \infty, \tag{1.21}$$

then every solution of (1.20) oscillates for every $r(t) > 0$, but one can easily see that this result cannot be applied when $p(t) = t^{-\alpha}$ for $\alpha > 1$.

Lin [38] considered the second-order nonlinear neutral differential equations

$$[x(t) - p(t)x(t-\tau)]'' + q(t)f(x(t-\sigma)) = 0, \quad t \geq 0, \tag{1.22}$$

where $0 \leq p(t) \leq 1,\ q(t) \geq 0,\ \tau,\ \sigma > 0$. The author investigated the oscillation for (1.22) when f is superlinear.

Wong [46, 47] studied the second-order neutral differential equations

$$[y(t) - py(t-\tau)]'' + q(t)f(y(t-\sigma)) = 0, \quad t \geq 0, \tag{1.23}$$

$q \in C[0,\infty),\ q(t) \geq 0,\ f \in C^1(-\infty,\infty),\ yf(y) > 0$ whenever $y \neq 0,\ f'(y) \geq 0$ for all y, and $0 < p < 1,\ \tau > 0,\ \sigma > 0$ are constants.

The main results for the oscillation of (1.23) in [46, 47] are the following.

Theorem 1.3 (see, [46, 47, Theorem 1]). *Suppose that f is superlinear. Then a solution of* (1.23) *is either oscillatory or tends to zero if and only if*

$$\int^{\infty} tq(t)\mathrm{d}t = \infty. \tag{1.24}$$

Theorem 1.4 (see, [46, 47, Theorem 2]). *Suppose that f is sublinear and in addition satisfies*

$$f(uv) \geq f(u)f(v), \quad uv \geq 0. \tag{1.25}$$

Then a solution of (1.23) *is either oscillatory or tends to zero if and only if*

$$\int^{\infty} f(t)q(t)\mathrm{d}t = \infty. \tag{1.26}$$

Li and Saker [40] investigated the second-order sublinear neutral delay difference equations

$$\Delta(a_n\Delta(x_n + p_n x_{n-\tau})) + q_n x_{n-\sigma}^{\gamma} = 0, \tag{1.27}$$

where $0 < \gamma < 1$ is a quotient of odd positive integers, $a_n > 0,\ \Delta a_n \geq 0, \sum_{n=0}^{\infty} 1/a_n = \infty,\ 0 \leq p_n < 1$, for all $n \geq 0$ and $q_n \geq 0$.

The main result for the oscillation of (1.27) in [40] is the following.

to Contents, page 1

Theorem 1.5 (see, [40, Theorem 2.1]). *Assume that there exists a positive sequence* $\{\rho_n\}$ *such that for every* $\alpha \geq 1$,

$$\limsup_{n\to\infty}\sum_{l=0}^{n}\left[\rho_l Q_l - \frac{a_{l-\sigma}(\alpha(l+1-\sigma))^{1-\gamma}(\Delta\rho_l)^2}{4\gamma\rho_l}\right] = \infty, \tag{1.28}$$

where $Q_n = q_n(1-p_{n-\sigma})^{\gamma}$. *Then every solution of* (1.27) *oscillates.*

Yildiz and Öcalan [41] studied the higher-order sublinear neutral delay difference equations of the type

$$\Delta^m(y_n + p_n y_{n-l}) + q_n y_{n-k}^{\alpha} = 0, \quad n \in \mathbb{N}, \tag{1.29}$$

where $0 < \alpha < 1$ is a ratio of odd positive integers. The authors established some oscillation criteria of (1.29).

The main results for the oscillation of (1.29) when $m = 2$ in [41] are the following.

Theorem 1.6 (see, [41, Theorem 2.1(a), $m = 2$]). *Assume that* $0 \leq p_n < 1$, *and*

$$\sum_{n=0}^{\infty} q_n\left[(1-p_{n-k})n\right]^{\alpha} = \infty. \tag{1.30}$$

Then all solutions of (1.29) *are oscillatory.*

Theorem 1.7 (see, [41, Theorem 2.2, $m = 2$]). *Assume that* $-1 < -p_2 \leq p_n \leq 0$, *where* $p_2 > 0$ *is a constant, and*

$$\sum_{n=0}^{\infty} q_n n^{\alpha} = \infty. \tag{1.31}$$

Then every solution of (1.29) *either oscillates or tends to zero as* $n \to \infty$.

Cheng [42] considered the oscillation of the second-order nonlinear neutral difference equations

$$\Delta\left(p_n(\Delta(x_n + c_n x_{n-\tau}))^{\gamma}\right) + q_n x_{n-\sigma}^{\beta} = 0 \tag{1.32}$$

and established some oscillation criteria of (1.32) by means of Riccati transformation techniques.

Following this trend, in this paper, we are concerned with oscillation of the second-order neutral delay dynamic equations of Emden-Fowler type

$$\left[a(t)(x(t) + r(t)x(\tau(t)))^{\Delta}\right]^{\Delta} + p(t)x^{\gamma}(\delta(t)) = 0, \quad t \in \mathbb{T}. \tag{1.33}$$

As we are interested in oscillatory behavior, we assume throughout this paper that the given time scales $\mathbb{T}$ are unbounded above; that is, it is a time scale interval of the form $[t_0, \infty)$ with $t_0 \in \mathbb{T}$.

We assume that $\gamma > 0$ is a quotient of odd positive integers, the delay functions τ and δ satisfy $\tau : \mathbb{T} \to \mathbb{T}$, $\delta : \mathbb{T} \to \mathbb{T}$, $\tau(t) \le t$, $\delta(t) \le t$ for all $t \in \mathbb{T}$, and $\lim_{t\to\infty}\tau(t) = \lim_{t\to\infty}\delta(t) = \infty$; a, r and p are real-valued rd-continuous functions defined on $\mathbb{T}$, $a(t) > 0$, $p(t) > 0$, $\int_{t_0}^{\infty} \Delta t/a(t) = \infty$.

We note that if $\mathbb{T} = \mathbb{R}$, then $\sigma(t) = t$, $\mu(t) = 0$, $x^{\Delta}(t) = x'(t)$, and (1.33) becomes the second-order nonlinear delay differential equation

$$\left[a(t)(x(t) + r(t)x(\tau(t)))'\right]' + p(t)x^{\gamma}(\delta(t)) = 0, \quad t \in \mathbb{R}. \tag{1.34}$$

If $\mathbb{T} = \mathbb{Z}$, then $\sigma(t) = t+1$, $\mu(t) = 1$, $x^{\Delta}(t) = \Delta x(t) = x(t+1) - x(t)$, and (1.33) becomes the second-order nonlinear delay differential equation

$$\Delta[a(t)\Delta(x(t) + r(t)x(\tau(t)))] + p(t)x^{\gamma}(\delta(t)) = 0, \quad t \in \mathbb{Z}. \tag{1.35}$$

In the case of $\gamma > 1$, (1.33) is the prototype of a wide class of nonlinear dynamic equations called Emden-Fowler sublinear dynamic equations, and if $\gamma < 1$, (1.33) is the prototype of dynamic equations called Emden-Fowler sublinear dynamic equations. It is interesting to study (1.33) because the continuous version, that is, (1.34), has several physical applications; see, for example, [1, 39], and when t is a discrete variable, it is (1.35), and it is also important in applications.

2. Main Results

In this section, we give some new oscillation criteria of (1.33). In order to prove our main results, we will use the formula

$$\left((x(t))^{\gamma}\right)^{\Delta} = \gamma \int_0^1 [hx^{\sigma}(t) + (1-h)x(t)]^{\gamma-1} x^{\Delta}(t)\mathrm{d}h, \tag{2.1}$$

which is a simple consequence of Keller's chain rule [3, Theorem 1.90]. Also, we need the following auxiliary results.

For the sake of convenience, we assume that

$$z(t) = x(t) + r(t)x(\tau(t)), \quad R(t,t_*) = a(t)\int_{t_*}^{t} \frac{\Delta s}{a(s)}, \quad \alpha(t,t_*) = \frac{\int_{t_*}^{\delta(t)} \Delta s/a(s)}{\int_{t_*}^{t} \Delta s/a(s)}, \quad t_* \ge t_0. \tag{2.2}$$

Lemma 2.1. *Assume that* (1.11) *holds,* $a^{\Delta}(t) \ge 0$, *and* $0 \le r(t) < 1$. *Then an eventually positive solution* x *of* (1.33) *eventually satisfies that*

$$z(t) \ge tz^{\Delta}(t) > 0, \quad z^{\Delta\Delta}(t) < 0, \quad \left(a(t)z^{\Delta}(t)\right)^{\Delta} < 0, \quad \frac{z(t)}{t} \text{ is nonincreasing.} \tag{2.3}$$

to Contents, page 1

Proof. From (1.11), the proof is similar to that of Saker et al. [24, Lemma 2.1], so it is omitted. □

Lemma 2.2. *Assume that*

$$\int_{t_0}^{\infty} p(t)\delta^{\gamma}(t)\Delta t = \infty, \tag{2.4}$$

$a^{\Delta}(t) \geq 0$, $-1 < -r_0 \leq r(t) \leq 0$, *and* $\lim_{t\to\infty} r(t) = r_1 > -1$. *Then an eventually positive solution* x *of* (1.33) *eventually satisfies that*

$$z(t) \geq t z^{\Delta}(t) > 0, \quad z^{\Delta\Delta}(t) < 0, \quad \left(a(t)z^{\Delta}(t)\right)^{\Delta} < 0, \quad \frac{z(t)}{t} \quad \text{is nonincreasing}, \tag{2.5}$$

or $\lim_{t\to\infty} x(t) = 0$.

Proof. Let x be an eventually positive solution of (1.33). Then there exists $t_1 \geq t_0$ such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\delta(t)) > 0$ for all $t \geq t_1$. Assume that $\lim_{t\to\infty} x(t) \neq 0$, that is, $\limsup_{t\to\infty} x(t) > 0$. Then, we have to show that (2.5) holds. It follows from (1.33) that

$$\left(a(t)z^{\Delta}(t)\right)^{\Delta} = -p(t)x^{\gamma}(\delta(t)) < 0, \quad t \geq t_1, \tag{2.6}$$

which implies that az^{Δ} is nonincreasing on $[t_1, \infty)_{\mathbb{T}}$. Since the function a is nondecreasing, z^{Δ} must be nonincreasing on $[t_1, \infty)_{\mathbb{T}}$, that is, z^{Δ} is eventually either positive or negative. In both cases, z is eventually monotonic, so that z has a limit at infinity (finite or infinite). This implies that $\lim_{t\to\infty} z(t) \neq 0$; that is, z is eventually positive (see [19, Lemma 3]). Then we proceed as in the proof of [24, Lemma 2.1] to obtain (2.5). The proof is complete. □

Lemma 2.3. *Assume that* $0 \leq r(t) < 1$. *Further,* x *is an eventually positive solution of* (1.33). *Then there exists a* $t_* \geq t_0$ *such that for* $t \geq t_*$,

$$z^{\Delta}(t) > 0, \quad \left(a(t)z^{\Delta}(t)\right)^{\Delta} < 0, \quad z(t) \geq R(t, t_*)z^{\Delta}(t), \quad z(\delta(t)) \geq \alpha(t, t_*)z(t). \tag{2.7}$$

Proof. Let x be an eventually positive solution of (1.33). Then there exists $t_1 \geq t_0$ such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\delta(t)) > 0$ for all $t \geq t_1$. It follows from (1.33) that (2.6) holds. From (2.6), we know that $a(t)z^{\Delta}(t)$ is an eventually decreasing function. We claim that $z^{\Delta}(t) > 0$ eventually. Otherwise, if there exists a $t_2 \geq t_1$ such that $z^{\Delta}(t) < 0$, by (2.6), we have

$$a(t)z^{\Delta}(t) \leq a(t_2)z^{\Delta}(t_2) = b < 0, \quad t \geq t_2. \tag{2.8}$$

Thus

$$z^{\Delta}(t) \leq b\frac{1}{a(t)}. \tag{2.9}$$

to Contents, page 1

Integrating the above inequality from t_2 to t leads to $\lim_{t\to\infty} z(t) = -\infty$, which contradicts $z(t) > 0$. Hence, $z^{\Delta}(t) > 0$ on $[t_2, \infty)_{\mathbb{T}}$. Therefore,

$$z(t) > z(t) - z(t_2) = \int_{t_2}^{t} \frac{a(s)z^{\Delta}(s)}{a(s)} \Delta s \geq \left(a(t)z^{\Delta}(t)\right) \int_{t_2}^{t} \frac{\Delta s}{a(s)}, \tag{2.10}$$

which yields

$$z(t) \geq \left(a(t) \int_{t_2}^{t} \frac{\Delta s}{a(s)}\right) z^{\Delta}(t). \tag{2.11}$$

Since $a(t)z^{\Delta}(t)$ is strictly decreasing, we have

$$z(t) - z(\delta(t)) = \int_{\delta(t)}^{t} \frac{a(s)z^{\Delta}(s)}{a(s)} \Delta s \leq a(\delta(t))z^{\Delta}(\delta(t)) \int_{\delta(t)}^{t} \frac{\Delta s}{a(s)}, \tag{2.12}$$

and so

$$\frac{z(t)}{z(\delta(t))} \leq 1 + \frac{a(\delta(t))z^{\Delta}(\delta(t))}{z(\delta(t))} \int_{\delta(t)}^{t} \frac{\Delta s}{a(s)}. \tag{2.13}$$

Also, we have that for large t,

$$z(\delta(t)) \geq z(\delta(t)) - z(t_2) = \int_{t_2}^{\delta(t)} \frac{a(s)z^{\Delta}(s)}{a(s)} \Delta s \geq a(\delta(t))z^{\Delta}(\delta(t)) \int_{t_2}^{\delta(t)} \frac{\Delta s}{a(s)}, \tag{2.14}$$

so we obtain

$$\frac{a(\delta(t))z^{\Delta}(\delta(t))}{z(\delta(t))} \leq \left(\int_{t_2}^{\delta(t)} \frac{\Delta s}{a(s)}\right)^{-1}. \tag{2.15}$$

Therefore, from (2.13), we have

$$z(\delta(t)) \geq \alpha(t, t_2) z(t). \tag{2.16}$$

This completes the proof. □

Lemma 2.4. *Assume that* $-1 < -r_0 \leq r(t) \leq 0$, $\lim_{t\to\infty} r(t) = r_1 > -1$. *Then an eventually positive solution* x *of* (1.33) *satisfies that, for sufficiently large* $t_* \geq t_0$,

$$z^{\Delta}(t) > 0, \quad \left(a(t)z^{\Delta}(t)\right)^{\Delta} < 0, \quad z(t) \geq R(t, t_*) z^{\Delta}(t), \quad z(\delta(t)) \geq \alpha(t, t_*) z(t), \quad t \geq t_*, \tag{2.17}$$

or $\lim_{t\to\infty} x(t) = 0$.

to Contents, page 1

Proof. The proof is similar to that of the proof Lemmas 2.2 and 2.3, so we omit the details. □

Theorem 2.5. *Assume that* (1.11) *holds,* $a^{\Delta}(t) \geq 0$, *and* $0 \leq r(t) < 1$. *Then every solution of* (1.33) *oscillates if the inequality*

$$y^{\Delta}(t) + A(t) y^{\gamma}(\delta(t)) \leq 0, \tag{2.18}$$

where

$$A(t) = p(t)(1 - r(\delta(t)))^{\gamma} \frac{(\delta(t))^{\gamma}}{(a(\delta(t)))^{\gamma}}, \tag{2.19}$$

has no eventually positive solution.

Proof. Suppose to the contrary that (1.33) has a nonoscillatory solution x. We may assume without loss of generality that there exists $t_1 \geq t_0$ such that $x(t) > 0$, $x(\tau(t)) > 0$ and $x(\delta(t)) > 0$ for all $t \geq t_1$. From Lemma 2.1, there is some $t_2 \geq t_1$ such that

$$x(t) = z(t) - r(t)x(\tau(t)) \geq z(t) - r(t)z(\tau(t)) \geq (1 - r(t))z(t), \quad t \geq t_2. \tag{2.20}$$

From (1.33), there exists a $t_3 \geq t_2$ such that

$$\left(a(t) z^{\Delta}(t)\right)^{\Delta} + p(t)(1 - r(\delta(t)))^{\gamma} (z(\delta(t)))^{\gamma} \leq 0, \quad t \geq t_3. \tag{2.21}$$

By Lemma 2.1, there exists a $t_4 \geq t_3$ such that

$$z(\delta(t)) \geq \delta(t) z^{\Delta}(\delta(t)). \tag{2.22}$$

Substituting the last inequality in (2.21) we obtain for $t \geq t_4$ that

$$\left(a(t) z^{\Delta}(t)\right)^{\Delta} + p(t)(1 - r(\delta(t)))^{\gamma} (\delta(t))^{\gamma} \left(z^{\Delta}(\delta(t))\right)^{\gamma} \leq 0. \tag{2.23}$$

Set $y(t) = a(t) z^{\Delta}(t)$. Then from (2.23), y is positive and satisfies the inequality (2.18), and this contradicts the assumption of our theorem. Thus every solution of (1.33) oscillates. This completes the proof. □

By [41, Lemma 1.1] and Theorem 2.5 in this paper, we have the following result.

Corollary 2.6. *If* $\mathbb{T} = \mathbb{Z}$, $a(t) = 1$, $\delta(t) = t - l$, l *is a positive integer, and* $0 \leq r(t) < 1$, *then every solution of* (1.33) *oscillates if*

$$\sum_{t=n_0}^{\infty} t^{\gamma} p(t)(1 - r(\delta(t)))^{\gamma} = \infty. \tag{2.24}$$

to Contents, page 1

Theorem 2.7. *Assume that* (2.4) *holds, and* $a^{\Delta}(t) \geq 0$, $-1 < -r_0 \leq r(t) \leq 0$, *and* $\lim_{t\to\infty} r(t) = r_1 > -1$. *Then every solution of* (1.33) *either oscillates or tends to zero as* $t \to \infty$ *if the inequality*

$$y^{\Delta}(t) + B(t) y^{\gamma}(\delta(t)) \leq 0, \tag{2.25}$$

where

$$B(t) = p(t) \frac{(\delta(t))^{\gamma}}{(a(\delta(t)))^{\gamma}}, \tag{2.26}$$

has no eventually positive solution.

Proof. Suppose to the contrary that (1.33) has a nonoscillatory solution x. We may assume without loss of generality that there exists $t_1 \geq t_0$ such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\delta(t)) > 0$ for all $t \geq t_1$.

From Lemma 2.2, if (i) holds, there is some $t_2 \geq t_1$ such that

$$x(t) = z(t) - r(t) x(\tau(t)) \geq z(t) > 0, \quad t \geq t_2. \tag{2.27}$$

From (1.33), there exists a $t_3 \geq t_2$ such that

$$\left(a(t) z^{\Delta}(t)\right)^{\Delta} + p(t)(z(\delta(t)))^{\gamma} \leq 0, \quad t \geq t_3. \tag{2.28}$$

By Lemma 2.2, there exists a $t_3 \geq t_2$ such that

$$z(\delta(t)) \geq \delta(t) z^{\Delta}(\delta(t)). \tag{2.29}$$

Substituting the last inequality in (2.28), we obtain for $t \geq t_3$ that

$$\left(a(t) z^{\Delta}(t)\right)^{\Delta} + p(t)(\delta(t))^{\gamma} \left(z^{\Delta}(\delta(t))\right)^{\gamma} \leq 0. \tag{2.30}$$

Set $y(t) = a(t) z^{\Delta}(t)$. Then from (2.30), y is positive and satisfies the inequality (2.25), and this contradicts the assumption of our theorem.

If (ii) holds, by Lemma 2.2, we have $\lim_{t\to\infty} x(t) = 0$. This completes the proof. □

By [41, Lemma 1.1] and Theorem 2.7 in this paper, we have the following result.

Corollary 2.8. *Assume that* $\mathbb{T} = \mathbb{Z}$, $a(t) = 1$, $\delta(t) = t - l$, *l is a positive integer,* $-1 < -r_0 \leq r(t) \leq 0$, *and* $\lim_{t\to\infty} r(t) = r > -1$. *Then every solution of* (1.33) *either oscillates or tends to zero as* $t \to \infty$ *if*

$$\sum_{t=n_0}^{\infty} t^{\gamma} p(t) = \infty. \tag{2.31}$$

Remark 2.9. Theorems 2.5 and 2.7 reduce the question of (1.33) to the absence of eventually positive solution (the oscillatory) of the differential inequalities (2.18) and (2.25).

to Contents, page 1

Remark 2.10. From Theorem 2.5, Theorem 2.7, and the results given in [7–9, 12, 14], we can obtain some oscillation criteria for (1.33) in the case when $\gamma = 1$, $a^{\Delta}(t) \geq 0$.

Theorem 2.11. *Assume that* (1.11) *holds,* $\gamma < 1$, $a^{\Delta}(t) \geq 0$, *and* $0 \leq r(t) < 1$. *Then every solution of* (1.33) *oscillates if*

$$\int_{t_0}^{\infty} \frac{p(s)}{(a(\delta(s)))^{\gamma}} (1 - r(\delta(s)))^{\gamma} (\delta(s))^{\gamma} \Delta s = \infty. \tag{2.32}$$

Proof. We assume that (1.33) has a nonoscillatory solution such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\delta(t)) > 0$ for all $t \geq t_1 \geq t_0$. By proceeding as in the proof of Theorem 2.5, we get (2.21). By Lemma 2.1, note that $(a(t)z^{\Delta}(t))^{\Delta} < 0$, and from Keller's chain rule, we obtain

$$\begin{aligned}
\left(\left(a(t)z^{\Delta}(t)\right)^{1-\gamma}\right)^{\Delta} &= (1-\gamma) \int_0^1 \left[h\left(a(t)z^{\Delta}(t)\right)^{\sigma} + (1-h)a(t)z^{\Delta}(t)\right]^{-\gamma} \left(a(t)z^{\Delta}(t)\right)^{\Delta} \mathrm{d}h \\
&\leq (1-\gamma) \int_0^1 \left[ha(t)z^{\Delta}(t) + (1-h)a(t)z^{\Delta}(t)\right]^{-\gamma} \left(a(t)z^{\Delta}(t)\right)^{\Delta} \mathrm{d}h \\
&= (1-\gamma)\left(a(t)z^{\Delta}(t)\right)^{-\gamma} \left(a(t)z^{\Delta}(t)\right)^{\Delta} < 0,
\end{aligned} \tag{2.33}$$

so

$$\left(a(t)z^{\Delta}(t)\right)^{-\gamma} \left(a(t)z^{\Delta}(t)\right)^{\Delta} \geq \frac{\left((a(t)z^{\Delta}(t))^{1-\gamma}\right)^{\Delta}}{1-\gamma}. \tag{2.34}$$

Using (2.21), we have

$$\begin{aligned}
0 &\geq \frac{(a(t)z^{\Delta}(t))^{\Delta} + p(t)(1-r(\delta(t)))^{\gamma}(z(\delta(t)))^{\gamma}}{(a(t)z^{\Delta}(t))^{\gamma}} \\
&= \left(a(t)z^{\Delta}(t)\right)^{-\gamma} \left(a(t)z^{\Delta}(t)\right)^{\Delta} + p(t)(1-r(\delta(t)))^{\gamma} \left(\frac{z(\delta(t))}{a(t)z^{\Delta}(t)}\right)^{\gamma} \\
&\geq \frac{\left((a(t)z^{\Delta}(t))^{1-\gamma}\right)^{\Delta}}{1-\gamma} + \frac{p(t)}{(a(\delta(t)))^{\gamma}} (1-r(\delta(t)))^{\gamma}(\delta(t))^{\gamma}.
\end{aligned} \tag{2.35}$$

Hence,

$$\frac{p(t)}{(a(\delta(t)))^{\gamma}} (1-r(\delta(t)))^{\gamma}(\delta(t))^{\gamma} \leq \frac{\left((a(t)z^{\Delta}(t))^{1-\gamma}\right)^{\Delta}}{\gamma - 1}. \tag{2.36}$$

to Contents, page 1

Upon integration we arrive at

$$\int_{t_1}^{t} \frac{p(s)}{(a(\delta(s)))^{\gamma}}(1-r(\delta(s)))^{\gamma}(\delta(s))^{\gamma}\Delta s \leq \int_{t_1}^{t} \frac{\left((a(s)z^{\Delta}(s))^{1-\gamma}\right)^{\Delta}}{\gamma-1}\Delta s \leq \frac{(a(t_1)z^{\Delta}(t_1))^{1-\gamma}}{1-\gamma}. \quad (2.37)$$

This contradicts (2.32) and finishes the proof. □

Theorem 2.12. *Assume that* (2.4) *holds, and* $\gamma < 1$, $a^{\Delta}(t) \geq 0$, $-1 < -r_0 \leq r(t) \leq 0$, *and* $\lim_{t\to\infty} r(t) = r_1 > -1$. *Then every solution of* (1.33) *either oscillates or tends to zero as* $t \to \infty$ *if*

$$\int_{t_0}^{\infty} \frac{p(s)}{(a(\delta(s)))^{\gamma}}(\delta(s))^{\gamma}\Delta s = \infty. \quad (2.38)$$

Proof. By Lemma 2.2, the proof is similar to that of the proof of Theorem 2.11, so we omit the details. □

Theorem 2.13. *Assume that* $\gamma < 1$ *and* $0 \leq r(t) < 1$. *Then every solution of* (1.33) *oscillates if*

$$\int_{t_0}^{\infty} \frac{p(s)}{(a(\delta(s)))^{\gamma}}(1-r(\delta(s)))^{\gamma}(R(\delta(s),t_*))^{\gamma}\Delta s = \infty \quad (2.39)$$

holds for all sufficiently large t_*.

Proof. By Lemma 2.3, the proof is similar to that of the proof Theorem 2.11, so we omit the details. □

Theorem 2.14. *Assume that* $\gamma < 1$, $-1 < -r_0 \leq r(t) \leq 0$, *and* $\lim_{t\to\infty} r(t) = r_1 > -1$. *Then every solution of* (1.33) *either oscillates or tends to zero as* $t \to \infty$ *if*

$$\int_{t_0}^{\infty} \frac{p(s)}{(a(\delta(s)))^{\gamma}}(R(\delta(s),t_*))^{\gamma}\Delta s = \infty \quad (2.40)$$

holds for all sufficiently large t_*.

Proof. By using Lemma 2.4 and (2.28), the proof is similar to that of the proof of Theorem 2.11, so we omit the details. □

Theorem 2.15. *Assume that* (1.11) *holds,* $\gamma \geq 1$, $a^{\Delta}(t) \geq 0$, *and* $0 \leq r(t) < 1$. *Then every solution of* (1.33) *oscillates if*

$$\limsup_{t\to\infty}\left\{\frac{t}{a(t)}\int_{t}^{\infty} p(s)(1-r(\delta(s)))^{\gamma}\left(\frac{\delta(s)}{s}\right)^{\gamma}\Delta s\right\} = \infty. \quad (2.41)$$

Proof. Suppose to the contrary that (1.33) has a nonoscillatory solution x. We may assume without loss of generality that there exists $t_1 \geq t_0$ such that $x(t) > 0$, $x(\tau(t)) > 0$, and

to Contents, page 1

$x(\delta(t)) > 0$ for all $t \geq t_1$. By proceeding as in the proof of Theorem 2.5, we get (2.21). Thus from Lemma 2.1, we have for $T \geq t \geq t_1$,

$$\int_t^T p(s)(1-r(\delta(s)))^\gamma (z(\delta(s)))^\gamma \Delta s \leq -\int_t^T \left(a(s)z^\Delta(s)\right)^\Delta \Delta s = a(t)z^\Delta(t) - a(T)z^\Delta(T), \quad (2.42)$$

and hence

$$\int_t^T p(s)(1-r(\delta(s)))^\gamma (z(\delta(s)))^\gamma \Delta s \leq a(t)z^\Delta(t). \quad (2.43)$$

This and Lemma 2.1 provide, for sufficiently large $t \in \mathbb{T}$,

$$\begin{aligned} z(t) \geq tz^\Delta(t) &\geq \frac{t}{a(t)} \int_t^\infty p(s)(1-r(\delta(s)))^\gamma (z(\delta(s)))^\gamma \Delta s \\ &\geq \frac{t}{a(t)} \int_t^\infty p(s)(1-r(\delta(s)))^\gamma \left(\frac{\delta(s)}{s}\right)^\gamma z^\gamma(s) \Delta s \\ &\geq z^\gamma(t) \left\{ \frac{t}{a(t)} \int_t^\infty p(s)(1-r(\delta(s)))^\gamma \left(\frac{\delta(s)}{s}\right)^\gamma \Delta s \right\}. \end{aligned} \quad (2.44)$$

So

$$\left\{ \frac{t}{a(t)} \int_t^\infty p(s)(1-r(\delta(s)))^\gamma \left(\frac{\delta(s)}{s}\right)^\gamma \Delta s \right\} \leq \left(\frac{1}{z(t)}\right)^{\gamma-1}. \quad (2.45)$$

We note that $\gamma \geq 1$ and $z^\Delta(t) > 0$ imply

$$\left\{ \frac{t}{a(t)} \int_t^\infty p(s)(1-r(\delta(s)))^\gamma \left(\frac{\delta(s)}{s}\right)^\gamma \Delta s \right\} \leq \left(\frac{1}{z(t_1)}\right)^{\gamma-1}. \quad (2.46)$$

This contradicts (2.41) and completes the proof. □

Theorem 2.16. *Assume that* (2.4) *holds, and* $\gamma \geq 1$, $a^\Delta(t) \geq 0$, $-1 < -r_0 \leq r(t) \leq 0$, *and* $\lim_{t\to\infty} r(t) = r_1 > -1$. *Then every solution of* (1.33) *either oscillates or tends to zero as* $t \to \infty$ *if*

$$\limsup_{t\to\infty} \left\{ \frac{t}{a(t)} \int_t^\infty p(s) \left(\frac{\delta(s)}{s}\right)^\gamma \Delta s \right\} = \infty. \quad (2.47)$$

Proof. By using Lemma 2.2 and (2.28), the proof is similar to that of the proof of Theorem 2.15, so we omit the details. □

to Contents, page 1

Theorem 2.17. *Assume that* $\gamma \geq 1,\ 0 \leq r(t) < 1$. *Then every solution of* (1.33) *oscillates if*

$$\limsup_{t\to\infty}\left\{\frac{R(t,t_*)}{a(t)}\int_t^{\infty} p(s)(1-r(\delta(s)))^{\gamma}(\alpha(s,t_*))^{\gamma}\Delta s\right\}=\infty \tag{2.48}$$

holds for all sufficiently large t_*.

Proof. Suppose to the contrary that (1.33) has a nonoscillatory solution x. We may assume without loss of generality that there exists $t_1 \geq t_0$ such that $x(t) > 0,\ x(\tau(t)) > 0$, and $x(\delta(t)) > 0$ for all $t \geq t_1$. By proceeding as in the proof of Theorem 2.5, we obtain (2.21). Thus from Lemma 2.3, we have, for $T \geq t \geq t_1$,

$$\int_t^T p(s)(1-r(\delta(s)))^{\gamma}(z(\delta(s)))^{\gamma}\Delta s \leq -\int_t^T \left(a(s)z^{\Delta}(s)\right)^{\Delta}\Delta s = a(t)z^{\Delta}(t)-a(T)z^{\Delta}(T), \tag{2.49}$$

and hence

$$\int_t^T p(s)(1-r(\delta(s)))^{\gamma}(z(\delta(s)))^{\gamma}\Delta s \leq a(t)z^{\Delta}(t). \tag{2.50}$$

This and Lemma 2.3 provide, for sufficiently large $t \in \mathbb{T}$,

$$\begin{aligned} z(t) &\geq R(t,t_*)z^{\Delta}(t) \geq \frac{R(t,t_*)}{a(t)}\int_t^{\infty} p(s)(1-r(\delta(s)))^{\gamma}(z(\delta(s)))^{\gamma}\Delta s \\ &\geq \frac{R(t,t_*)}{a(t)}\int_t^{\infty} p(s)(1-r(\delta(s)))^{\gamma}(\alpha(s,t_*))^{\gamma}z^{\gamma}(s)\Delta s \\ &\geq z^{\gamma}(t)\left\{\frac{R(t,t_*)}{a(t)}\int_t^{\infty} p(s)(1-r(\delta(s)))^{\gamma}(\alpha(s,t_*))^{\gamma}\Delta s\right\}. \end{aligned} \tag{2.51}$$

So

$$\left\{\frac{R(t,t_*)}{a(t)}\int_t^{\infty} p(s)(1-r(\delta(s)))^{\gamma}(\alpha(s,t_*))^{\gamma}\Delta s\right\} \leq \left(\frac{1}{z(t)}\right)^{\gamma-1}. \tag{2.52}$$

We note that $\gamma \geq 1$ and $z^{\Delta}(t) > 0$ imply

$$\left\{\frac{R(t,t_*)}{a(t)}\int_t^{\infty} p(s)(1-r(\delta(s)))^{\gamma}(\alpha(s,t_*))^{\gamma}\Delta s\right\} \leq \left(\frac{1}{z(t_1)}\right)^{\gamma-1}. \tag{2.53}$$

This contradicts (2.48) and completes the proof. □

to Contents, page 1

Theorem 2.18. *Assume that* (2.4) *holds, and* $\gamma \geq 1$, $a^{\Delta}(t) \geq 0$, $-1 < -r_0 \leq r(t) \leq 0$, *and* $\lim_{t\to\infty} r(t) = r > -1$. *Then every solution of* (1.33) *either oscillates or tends to zero as* $t \to \infty$ *if*

$$\limsup_{t\to\infty}\left\{\frac{R(t,t_*)}{a(t)}\int_t^{\infty} p(s)(\alpha(s,t_*))^{\gamma}\Delta s\right\} = \infty \tag{2.54}$$

holds for all sufficiently large t_*.

Proof. By using Lemma 2.4 and (2.28), the proof is similar to that of the proof of Theorem 2.17, so we omit the details. □

Theorem 2.19. *Assume that* (1.11) *holds,* $\gamma > 1$, $a^{\Delta}(t) \geq 0$, *and* $0 \leq r(t) < 1$. *Then every solution of* (1.33) *oscillates if*

$$\int_{t_0}^{\infty} \sigma(s)\frac{p(s)}{a(s)}(1 - r(\delta(s)))^{\gamma}\left(\frac{\delta(s)}{\sigma(s)}\right)^{\gamma}\Delta s = \infty. \tag{2.55}$$

Proof. We assume that (1.33) has a nonoscillatory solution such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\delta(t)) > 0$ for all $t \geq t_1 \geq t_0$. By proceeding as in the proof of Theorem 2.5, we get (2.21). Define the function

$$\omega(t) = \frac{ta(t)z^{\Delta}(t)}{z^{\gamma}(t)}, \quad t \geq t_1. \tag{2.56}$$

By Lemma 2.1, $\omega(t) > 0$. We calculate

$$\omega^{\Delta}(t) = \left\{a(t)z^{\Delta}(t) + \sigma(t)\left(a(t)z^{\Delta}(t)\right)^{\Delta}\right\}\left(z^{-\gamma}(t)\right)^{\sigma} + ta(t)z^{\Delta}(t)\left(z^{-\gamma}(t)\right)^{\Delta}. \tag{2.57}$$

From (2.21), we have

$$\omega^{\Delta}(t) \leq a(t)z^{\Delta}(t)\left(z^{-\gamma}(t)\right)^{\sigma} - \sigma(t)p(t)(1 - r(\delta(t)))^{\gamma}\left(\frac{z(\delta(t))}{z(\sigma(t))}\right)^{\gamma} + ta(t)z^{\Delta}(t)\left(z^{-\gamma}(t)\right)^{\Delta}, \tag{2.58}$$

and by Lemma 2.1, we have

$$\omega^{\Delta}(t) \leq a(t)z^{\Delta}(t)\left(z^{-\gamma}(t)\right)^{\sigma} - \sigma(t)p(t)(1 - r(\delta(t)))^{\gamma}\left(\frac{\delta(t)}{\sigma(t)}\right)^{\gamma}, \tag{2.59}$$

because $(z^{-\gamma}(t))^{\Delta} \leq 0$ due to Keller's chain rule. Since

$$\begin{aligned}\left((z(t))^{1-\gamma}\right)^{\Delta} &= (1-\gamma)\int_0^1 [hz^{\sigma}(t) + (1-h)z(t)]^{-\gamma} z^{\Delta}(t)\mathrm{d}h \\ &\leq (1-\gamma)\int_0^1 [hz^{\sigma}(t) + (1-h)z^{\sigma}(t)]^{-\gamma} z^{\Delta}(t)\mathrm{d}h = (1-\gamma)(z^{\sigma}(t))^{-\gamma}z^{\Delta}(t),\end{aligned} \tag{2.60}$$

to Contents, page 1

thus

$$\omega^{\Delta}(t) \le a(t)\frac{\left((z(t))^{1-\gamma}\right)^{\Delta}}{1-\gamma} - \sigma(t)p(t)(1-r(\delta(t)))^{\gamma}\left(\frac{\delta(t)}{\sigma(t)}\right)^{\gamma}. \tag{2.61}$$

Upon integration we arrive at

$$\begin{aligned}
&\int_{t_1}^{t} \sigma(s)\frac{p(s)}{a(s)}(1-r(\delta(s)))^{\gamma}\left(\frac{\delta(s)}{\sigma(s)}\right)^{\gamma}\Delta s \\
&\quad \le \int_{t_1}^{t}\left\{\frac{\left((z(s))^{1-\gamma}\right)^{\Delta}}{1-\gamma} - \frac{\omega^{\Delta}(s)}{a(s)}\right\}\Delta s \\
&\quad = \frac{(z(t))^{1-\gamma}}{1-\gamma} - \frac{(z(t_1))^{1-\gamma}}{1-\gamma} - \int_{t_1}^{t}\frac{\omega^{\Delta}(s)}{a(s)}\Delta s \\
&\quad = \frac{(z(t))^{1-\gamma}}{1-\gamma} - \frac{(z(t_1))^{1-\gamma}}{1-\gamma} - \frac{\omega(t)}{a(t)} + \frac{\omega(t_1)}{a(t_1)} + \int_{t_1}^{t}\omega^{\sigma}(s)\left(\frac{1}{a(s)}\right)^{\Delta}\Delta s.
\end{aligned} \tag{2.62}$$

Noting that $(1/a(t))^{\Delta} \le 0$, we have

$$\int_{t_1}^{t} \sigma(s)\frac{p(s)}{a(s)}(1-r(\delta(s)))^{\gamma}\left(\frac{\delta(s)}{\sigma(s)}\right)^{\gamma}\Delta s \le \frac{(z(t_1))^{1-\gamma}}{\gamma-1} + \frac{\omega(t_1)}{a(t_1)}. \tag{2.63}$$

This contradicts (2.55) and finishes the proof. □

Theorem 2.20. *Assume that* (2.4) *holds, and* $\gamma > 1$, $a^{\Delta}(t) \ge 0$, $-1 < -r_0 \le r(t) \le 0$, *and* $\lim_{t\to\infty} r(t) = r_1 > -1$. *Then every solution of* (1.33) *either oscillates or tends to zero as* $t \to \infty$ *if*

$$\int_{t_0}^{\infty} \sigma(s)p(s)\left(\frac{\delta(s)}{\sigma(s)}\right)^{\gamma}\Delta s = \infty. \tag{2.64}$$

Proof. By using Lemma 2.2 and (2.28), the proof is similar to that of the proof of Theorem 2.19, so we omit the details. □

In the following, we use a Riccati transformation technique to establish new oscillation criteria for (1.33).

Theorem 2.21. *Assume that* $\gamma \ge 1$, *and* $0 \le r(t) < 1$. *Furthermore, suppose that there exists a positive* Δ*-differentiable function* η *such that for all sufficiently large* t_*, *and for all constants* $M > 0$,

to Contents, page 1

for $t_1 \geq t_*$,

$$\limsup_{t\to\infty} \int_{t_1}^{t} \left[\eta(s)p(s)(1-r(\delta(s)))^{\gamma}(\alpha(s,t_*))^{\gamma} - \frac{a(s)\left(\eta^{\Delta}(s)\right)^2}{4\gamma M^{\gamma-1}\eta(s)} \right] \Delta s = \infty. \tag{2.65}$$

Then every solution of (1.33) *oscillates.*

Proof. We assume that (1.33) has a nonoscillatory solution such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\delta(t)) > 0$ for all $t \geq t_1 \geq t_0$. By proceeding as in the proof of Theorem 2.5, we get (2.21). Define the function ω by the Riccati substitution

$$\omega(t) = \eta(t)\frac{a(t)z^{\Delta}(t)}{z^{\gamma}(t)}, \quad t \geq t_1. \tag{2.66}$$

Then $\omega(t) > 0$. By the product rule and then the quotient rule

$$\begin{aligned} \omega^{\Delta}(t) &= \left(a(t)z^{\Delta}(t)\right)^{\sigma}\left[\frac{\eta(t)}{z^{\gamma}(t)}\right]^{\Delta} + \frac{\eta(t)}{z^{\gamma}(t)}\left(a(t)z^{\Delta}(t)\right)^{\Delta} \\ &= \frac{\eta(t)}{z^{\gamma}(t)}\left(a(t)z^{\Delta}(t)\right)^{\Delta} + \left(a(t)z^{\Delta}(t)\right)^{\sigma}\left[\frac{z^{\gamma}(t)\eta^{\Delta}(t) - \eta(t)(z^{\gamma}(t))^{\Delta}}{z^{\gamma}(t)(z^{\sigma}(t))^{\gamma}}\right]. \end{aligned} \tag{2.67}$$

In view of (2.21) and (2.66), we have

$$\omega^{\Delta}(t) \leq -\eta(t)p(t)(1-r(\delta(t)))^{\gamma}\left(\frac{z(\delta(t))}{z(t)}\right)^{\gamma} + \frac{\eta^{\Delta}(t)}{\eta^{\sigma}(t)}\omega^{\sigma}(t) - \frac{\eta(t)\left(a(t)z^{\Delta}(t)\right)^{\sigma}(z^{\gamma}(t))^{\Delta}}{z^{\gamma}(t)(z^{\sigma}(t))^{\gamma}}. \tag{2.68}$$

By the chain rule and $\gamma \geq 1$, we obtain

$$(z^{\gamma}(t))^{\Delta} \geq \gamma z^{\gamma-1}(t)z^{\Delta}(t) \geq \gamma M^{\gamma-1}z^{\Delta}(t), \tag{2.69}$$

where $M = z(t_1) > 0$. In view of $(a(t)z^{\Delta}(t))^{\Delta} < 0$, we have

$$a(t)z^{\Delta}(t) \geq \left(a(t)z^{\Delta}(t)\right)^{\sigma}, \tag{2.70}$$

and by Lemma 2.3, we see that

$$\omega^{\Delta}(t) \leq -\eta(t)p(t)(1-r(\delta(t)))^{\gamma}(\alpha(t,t_*))^{\gamma} + \frac{\eta^{\Delta}(t)}{\eta^{\sigma}(t)}\omega^{\sigma}(t) - \frac{\gamma M^{\gamma-1}\eta(t)}{a(t)\left(\eta^{\sigma}(t)\right)^2}(\omega^{\sigma}(t))^2. \tag{2.71}$$

to Contents, page 1

Integrating (2.71) from t_1 to t, we obtain

$$\begin{aligned}
&\int_{t_1}^{t} \eta(s)p(s)(1-r(\delta(s)))^{\gamma}(\alpha(s,t_*))^{\gamma}\Delta s \\
&\quad\le -\int_{t_1}^{t} \omega^{\Delta}(s)\Delta s \\
&\qquad+\int_{t_1}^{t} \frac{\eta^{\Delta}(s)}{\eta^{\sigma}(s)}\omega^{\sigma}(s)\Delta s-\int_{t_1}^{t} \frac{\gamma M^{\gamma-1}\eta(s)}{a(s)\left(\eta^{\sigma}(s)\right)^{2}}(\omega^{\sigma}(s))^{2}\Delta s.
\end{aligned} \tag{2.72}$$

Hence

$$\int_{t_1}^{t}\left[\eta(s)p(s)(1-r(\delta(s)))^{\gamma}(\alpha(s,t_*))^{\gamma}-\frac{a(s)\left(\eta^{\Delta}(s)\right)^{2}}{4\gamma M^{\gamma-1}\eta(s)}\right]\Delta s\le\omega(t_1), \tag{2.73}$$

which contradicts condition (2.65). The proof is complete. □

Theorem 2.22. *Assume that* $\gamma\ge 1$, $-1<-r_0\le r(t)\le 0$, *and* $\lim_{t\to\infty}r(t)=r_1>-1$. *If there exists a positive* Δ*-differentiable function* η *such that for all sufficiently large* t_*, *and for all constants* $M>0$, *for* $t_1\ge t_*$,

$$\limsup_{t\to\infty}\int_{t_1}^{t}\left[\eta(s)p(s)(\alpha(s,t_*))^{\gamma}-\frac{a(s)\left(\eta^{\Delta}(s)\right)^{2}}{4\gamma M^{\gamma-1}\eta(s)}\right]\Delta s=\infty, \tag{2.74}$$

then every solution of (1.33) *either oscillates or tends to zero as* $t\to\infty$.

Proof. By Lemma 2.4 and (2.28), the proof is similar to that of the proof of Theorem 2.21, so we omit the details. □

Theorem 2.23. *Assume that* (1.11) *holds,* $\gamma\le 1$, $a^{\Delta}(t)\ge 0$, *and* $0\le r(t)<1$. *Furthermore, suppose that there exists a positive* Δ*-differentiable function* η *such that for all sufficiently large* t_1, *and for all constants* $M>0$,

$$\limsup_{t\to\infty}\int_{t_1}^{t}\left[\eta(s)p(s)(1-r(\delta(s)))^{\gamma}\left(\frac{\delta(s)}{s}\right)^{\gamma}-\frac{a(s)\left(\eta^{\Delta}(s)\right)^{2}}{4\gamma M^{\gamma-1}(\sigma(s))^{\gamma-1}\eta(s)}\right]\Delta s=\infty. \tag{2.75}$$

Then every solution of (1.33) *oscillates.*

Proof. We assume that (1.33) has a nonoscillatory solution such that $x(t)>0$, $x(\tau(t))>0$, and $x(\delta(t))>0$ for all $t\ge t_1\ge t_0$. By proceeding as in the proof of Theorem 2.5, we obtain (2.21).

to Contents, page 1

Define the function ω by the Riccati substitution as (2.66). Then $\omega(t) > 0$. By the product rule and then the quotient rule

$$\begin{aligned}\omega^{\Delta}(t) &= \left(a(t)z^{\Delta}(t)\right)^{\sigma}\left[\frac{\eta(t)}{z^{\gamma}(t)}\right]^{\Delta} + \frac{\eta(t)}{z^{\gamma}(t)}\left(a(t)z^{\Delta}(t)\right)^{\Delta} \\ &= \frac{\eta(t)}{z^{\gamma}(t)}\left(a(t)z^{\Delta}(t)\right)^{\Delta} + \left(a(t)z^{\Delta}(t)\right)^{\sigma}\left[\frac{z^{\gamma}(t)\eta^{\Delta}(t) - \eta(t)(z^{\gamma}(t))^{\Delta}}{z^{\gamma}(t)(z^{\sigma}(t))^{\gamma}}\right].\end{aligned} \tag{2.76}$$

In view of (2.21) and (2.66), we have

$$\omega^{\Delta}(t) \le -\eta(t)p(t)(1-r(\delta(t)))^{\gamma}\left(\frac{z(\delta(t))}{z(t)}\right)^{\gamma} + \frac{\eta^{\Delta}(t)}{\eta^{\sigma}(t)}\omega^{\sigma}(t) - \frac{\eta(t)\left(a(t)z^{\Delta}(t)\right)^{\sigma}(z^{\gamma}(t))^{\Delta}}{z^{\gamma}(t)(z^{\sigma}(t))^{\gamma}}. \tag{2.77}$$

From the chain rule and $\gamma \le 1$, we get

$$(z^{\gamma}(t))^{\Delta} \ge \gamma z^{\gamma-1}(\sigma(t))z^{\Delta}(t). \tag{2.78}$$

Noting that $z(t)/t$ is nonincreasing, and there exists a constant $M > 0$, such that $z(t) \le Mt$, hence we have

$$(z^{\gamma}(t))^{\Delta} \ge \gamma z^{\gamma-1}(\sigma(t))z^{\Delta}(t) \ge \gamma M^{\gamma-1}(\sigma(t))^{\gamma-1}z^{\Delta}(t). \tag{2.79}$$

In view of $(a(t)z^{\Delta}(t))^{\Delta} < 0$, we have

$$a(t)z^{\Delta}(t) \ge \left(a(t)z^{\Delta}(t)\right)^{\sigma}, \tag{2.80}$$

and by Lemma 2.1, we see that

$$\omega^{\Delta}(t) \le -\eta(t)p(t)(1-r(\delta(t)))^{\gamma}\left(\frac{\delta(t)}{t}\right)^{\gamma} + \frac{\eta^{\Delta}(t)}{\eta^{\sigma}(t)}\omega^{\sigma}(t) - \frac{\gamma M^{\gamma-1}(\sigma(t))^{\gamma-1}\eta(t)}{a(t)\left(\eta^{\sigma}(t)\right)^{2}}(\omega^{\sigma}(t))^{2}. \tag{2.81}$$

Integrating (2.81) from t_1 to t, we obtain

$$\begin{aligned}&\int_{t_1}^{t}\eta(s)p(s)(1-r(\delta(s)))^{\gamma}\left(\frac{\delta(s)}{s}\right)^{\gamma}\Delta s \\ &\quad\le -\int_{t_1}^{t}\omega^{\Delta}(s)\Delta s + \int_{t_1}^{t}\frac{\eta^{\Delta}(s)}{\eta^{\sigma}(s)}\omega^{\sigma}(s)\Delta s - \int_{t_1}^{t}\frac{\gamma M^{\gamma-1}(\sigma(s))^{\gamma-1}\eta(s)}{a(s)\left(\eta^{\sigma}(s)\right)^{2}}(\omega^{\sigma}(s))^{2}\Delta s.\end{aligned} \tag{2.82}$$

to Contents, page 1

Hence

$$\int_{t_1}^{t}\left[\eta(s)p(s)(1-r(\delta(s)))^{\gamma}\left(\frac{\delta(s)}{s}\right)^{\gamma}-\frac{a(s)\left(\eta^{\Delta}(s)\right)^{2}}{4\gamma M^{\gamma-1}(\sigma(s))^{\gamma-1}\eta(s)}\right]\Delta s\leq\omega(t_1), \tag{2.83}$$

which contradicts condition (2.75). The proof is complete. □

Theorem 2.24. *Assume that* (2.4) *holds,* $\gamma \leq 1$, $a^{\Delta}(t) \geq 0$, $-1 < -r_0 \leq r(t) \leq 0$, *and* $\lim_{t\to\infty} r(t) = r_1 > -1$. *If there exists a positive* Δ*-differentiable function* η *such that for all sufficiently large* t_1, *and for all constants* $M > 0$,

$$\limsup_{t\to\infty}\int_{t_1}^{t}\left[\eta(s)p(s)\left(\frac{\delta(s)}{s}\right)^{\gamma}-\frac{a(s)\left(\eta^{\Delta}(s)\right)^{2}}{4\gamma M^{\gamma-1}(\sigma(s))^{\gamma-1}\eta(s)}\right]\Delta s=\infty, \tag{2.84}$$

then every solution of (1.33) *either oscillates or tends to zero as* $t \to \infty$.

Proof. By Lemma 2.2 and (2.28), the proof is similar to that of the proof of Theorem 2.23, so we omit the details. □

Theorem 2.25. *Assume that* $\gamma \leq 1$, $a^{\Delta}(t) \leq 0$, *and* $0 \leq r(t) < 1$. *Furthermore, suppose that there exists a positive* Δ*-differentiable function* η *such that for all sufficiently large* t_*, *and for all constants* $M > 0$, *for* $t_1 \geq t_*$,

$$\limsup_{t\to\infty}\int_{t_1}^{t}\left[\eta(s)p(s)(1-r(\delta(s)))^{\gamma}(\alpha(s,t_*))^{\gamma}-\frac{a(s)(\sigma(s))^{1-\gamma}\left(\eta^{\Delta}(s)\right)^{2}}{4\gamma M^{\gamma-1}(a(\sigma(s)))^{1-\gamma}\eta(s)}\right]\Delta s=\infty. \tag{2.85}$$

Then every solution of (1.33) *oscillates.*

Proof. We assume that (1.33) has a nonoscillatory solution such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\delta(t)) > 0$ for all $t \geq t_1 \geq t_0$. By proceeding as in the proof of Theorem 2.5, we have (2.21). Define the function ω by the Riccati substitution as (2.66). Then $\omega(t) > 0$. By the product rule and then the quotient rule

$$\begin{aligned}\omega^{\Delta}(t)&=\left(a(t)z^{\Delta}(t)\right)^{\sigma}\left[\frac{\eta(t)}{z^{\gamma}(t)}\right]^{\Delta}+\frac{\eta(t)}{z^{\gamma}(t)}\left(a(t)z^{\Delta}(t)\right)^{\Delta}\\&=\frac{\eta(t)}{z^{\gamma}(t)}\left(a(t)z^{\Delta}(t)\right)^{\Delta}+\left(a(t)z^{\Delta}(t)\right)^{\sigma}\left[\frac{z^{\gamma}(t)\eta^{\Delta}(t)-\eta(t)(z^{\gamma}(t))^{\Delta}}{z^{\gamma}(t)(z^{\sigma}(t))^{\gamma}}\right].\end{aligned} \tag{2.86}$$

In view of (2.21) and (2.66), we have

$$\omega^{\Delta}(t)\leq-\eta(t)p(t)(1-r(\delta(t)))^{\gamma}\left(\frac{z(\delta(t))}{z(t)}\right)^{\gamma}+\frac{\eta^{\Delta}(t)}{\eta^{\sigma}(t)}\omega^{\sigma}(t)-\frac{\eta(t)\left(a(t)z^{\Delta}(t)\right)^{\sigma}(z^{\gamma}(t))^{\Delta}}{z^{\gamma}(t)(z^{\sigma}(t))^{\gamma}}. \tag{2.87}$$

to Contents, page 1

By the chain rule and $\gamma \le 1$, we obtain

$$(z^{\gamma}(t))^{\Delta} \ge \gamma z^{\gamma-1}(\sigma(t))z^{\Delta}(t), \tag{2.88}$$

and noting that $(a(t)z^{\Delta}(t))^{\Delta} < 0$ and there exists a constant $L > 0$ such that $a(t)z^{\Delta}(t) \le L$, so

$$z(t) = z(t_1) + \int_{t_1}^{t} z^{\Delta}(s)\Delta s \le z(t_1) + \int_{t_1}^{t} \frac{L}{a(s)}\Delta s. \tag{2.89}$$

From $a^{\Delta}(t) \le 0$, there exists a positive constant M such that

$$z(t) \le z(t_1) + \frac{L}{a(t)}(t - t_1) = \frac{z(t_1)a(t) + L(t - t_1)}{a(t)} \le \frac{Mt}{a(t)}. \tag{2.90}$$

Hence

$$(z^{\gamma}(t))^{\Delta} \ge \gamma z^{\gamma-1}(\sigma(t))z^{\Delta}(t) \ge \gamma M^{\gamma-1}\left(\frac{\sigma(t)}{a(\sigma(t))}\right)^{\gamma-1} z^{\Delta}(t). \tag{2.91}$$

In view of $(a(t)z^{\Delta}(t))^{\Delta} < 0$, we have

$$a(t)z^{\Delta}(t) \ge \left(a(t)z^{\Delta}(t)\right)^{\sigma}, \tag{2.92}$$

and by Lemma 2.3, we see that

$$\begin{aligned} \omega^{\Delta}(t) \le{}& -\eta(t)p(t)(1 - r(\delta(t)))^{\gamma}(\alpha(t,t_*))^{\gamma} \\ &+ \frac{\eta^{\Delta}(t)}{\eta^{\sigma}(t)}\omega^{\sigma}(t) - \frac{\gamma M^{\gamma-1}\eta(t)}{a(t)(\eta^{\sigma}(t))^2}\left(\frac{\sigma(t)}{a(\sigma(t))}\right)^{\gamma-1}(\omega^{\sigma}(t))^2. \end{aligned} \tag{2.93}$$

Integrating (2.93) from t_1 to t, we obtain

$$\begin{aligned} &\int_{t_1}^{t} \eta(s)p(s)(1 - r(\delta(s)))^{\gamma}(\alpha(s,t_*))^{\gamma}\Delta s \\ &\le -\int_{t_1}^{t} \omega^{\Delta}(s)\Delta s + \int_{t_1}^{t} \frac{\eta^{\Delta}(s)}{\eta^{\sigma}(s)}\omega^{\sigma}(s)\Delta s - \int_{t_1}^{t} \frac{\gamma M^{\gamma-1}\eta(s)}{a(s)(\eta^{\sigma}(s))^2}\left(\frac{\sigma(s)}{a(\sigma(s))}\right)^{\gamma-1}(\omega^{\sigma}(s))^2\Delta s. \end{aligned} \tag{2.94}$$

Thus

$$\int_{t_1}^{t}\left[\eta(s)p(s)(1 - r(\delta(s)))^{\gamma}(\alpha(s,t_*))^{\gamma} - \frac{a(s)(\sigma(s))^{1-\gamma}(\eta^{\Delta}(s))^2}{4\gamma M^{\gamma-1}(a(\sigma(s)))^{1-\gamma}\eta(s)}\right]\Delta s \le \omega(t_1), \tag{2.95}$$

which contradicts condition (2.85). The proof is complete. □

to Contents, page 1

Theorem 2.26. *Assume that* $\gamma \leq 1$, $a^{\Delta}(t) \leq 0$, $-1 < -r_0 \leq r(t) \leq 0$, *and* $\lim_{t\to\infty} r(t) = r > -1$. *If there exists a positive* Δ*-differentiable function* η *such that for all sufficiently large* t_*, *and for all constants* $M > 0$, *for* $t_1 \geq t_*$,

$$\limsup_{t\to\infty} \int_{t_1}^{t} \left[\eta(s)p(s)(\alpha(s,t_*))^{\gamma} - \frac{a(s)(\sigma(s))^{1-\gamma}(\eta^{\Delta}(s))^2}{4\gamma M^{\gamma-1}(a(\sigma(s)))^{1-\gamma}\eta(s)} \right] \Delta s = \infty, \tag{2.96}$$

then every solution of (1.33) *either oscillates or tends to zero as* $t \to \infty$.

Proof. By Lemma 2.4 and (2.28), the proof is similar to that of the proof of Theorem 2.25, so we omit the details. □

3. Conclusions

In this paper, we consider the oscillation of second-order Emden-Fowler neutral delay dynamic equations (1.33). In some sense, our results extend and improve the results in [7, 32, 34, 35, 40, 41]. For example, Theorems 2.5, 2.11, 2.13, and 2.23 give some answers for the open problem posed by [34] since these results can be applied to (1.33) when $\gamma < 1$, Theorems 2.7, 2.12, 2.14, 2.16, 2.18, 2.20, 2.22, 2.24, and 2.26 correct an error in [35]. Theorem 2.15 includes the results of [7, Theorem 4.4], [32, Theorem 3.1], Theorem 2.11 includes the result of [32, Theorem 3.5], Theorem 2.11 and Corollary 2.6 include the result of [41, Theorem 2.1(a), $m = 2$], Corollary 2.8 includes result of [41, Theorem 2.2, $m = 2$], Theorem 2.13 does not require the conditions $a^{\Delta}(t) \geq 0$, so it improves the results of [40], and Theorems 2.17 and 2.21 improve the results in [34] since these results can be applied when $a^{\Delta}(t) \leq 0$.

The main results in this paper require that $\int_{t_0}^{\infty} \Delta t / a(t) = \infty$; it would be interesting to find another method to study (1.33) when $\int_{t_0}^{\infty} \Delta t / a(t) < \infty$. Additional examples may also be given; due to the limited space, we leave this to the interested reader.

Acknowledgments

The authors sincerely thank the reviewers for their valuable suggestions and useful comments that have lead to the present improved version of the original manuscript. This research is supported by the Natural Science Foundation of China (11071143, 60904024), China Postdoctoral Science Foundation funded project (20080441126, 200902564), Shandong Postdoctoral funded project (200802018) and supported by the Natural Science Foundation of Shandong (Y2008A28, ZR2009AL003), and by University of Jinan Research Funds for Doctors (XBS0843).

References

[1] S. Hilger, "Analysis on measure chains—a unified approach to continuous and discrete calculus," *Results in Mathematics*, vol. 18, no. 1-2, pp. 18–56, 1990.
[2] V. Kac and P. Cheung, *Quantum Calculus*, Universitext, Springer, New York, NY, USA, 2002.
[3] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales*, Birkhäauser, Boston, Mass, USA, 2001.
[4] V. Spedding, *Taming Nature's Numbers*, New Scientist, 2003.
[5] R. P. Agarwal, M. Bohner, D. O'Regan, and A. Peterson, "Dynamic equations on time scales: a survey," *Journal of Computational and Applied Mathematics*, vol. 141, no. 1-2, pp. 1–26, 2002.

to Contents, page 1

[6] M. Bohner and A. Peterson, *Advances in Dynamic Equations on Time Scales*, Birkhäauser, Boston, Mass, USA, 2003.

[7] R. P. Agarwal, M. Bohner, and S. H. Saker, "Oscillation of second order delay dynamic equations," *The Canadian Applied Mathematics Quarterly*, vol. 13, no. 1, pp. 1–17, 2005.

[8] R. P. Agarwal and M. Bohner, "An oscillation criterion for first order delay dynamic equations," *Functional Differential Equations*, vol. 16, no. 1, pp. 11–17, 2009.

[9] M. Bohner, B. Karpuz, and Ö. Öcalan, "Iterated oscillation criteria for delay dynamic equations of first order," *Advances in Difference Equations*, vol. 2008, Article ID 458687, 12 pages, 2008.

[10] M. Bohner and S. H. Saker, "Oscillation of second order nonlinear dynamic equations on time scales," *The Rocky Mountain Journal of Mathematics*, vol. 34, no. 4, pp. 1239–1254, 2004.

[11] Y. Şahiner, "Oscillation of second-order delay differential equations on time scales," *Nonlinear Analysis*, vol. 63, no. 5–7, pp. e1073–e1080, 2005.

[12] Y. Şahiner and I. P. Stavroulakis, "Oscillations of first order delay dynamic equations," *Dynamic Systems and Applications*, vol. 15, no. 3-4, pp. 645–656, 2006.

[13] B. G. Zhang and S. Zhu, "Oscillation of second-order nonlinear delay dynamic equations on time scales," *Computers & Mathematics with Applications*, vol. 49, no. 4, pp. 599–609, 2005.

[14] B. G. Zhang and X. Deng, "Oscillation of delay differential equations on time scales," *Mathematical and Computer Modelling*, vol. 36, no. 11–13, pp. 1307–1318, 2002.

[15] L. Erbe, A. Peterson, and S. H. Saker, "Oscillation criteria for second-order nonlinear dynamic equations on time scales," *Journal of the London Mathematical Society*, vol. 67, no. 3, pp. 701–714, 2003.

[16] L. Erbe, A. Peterson, and S. H. Saker, "Oscillation criteria for second-order nonlinear delay dynamic equations," *Journal of Mathematical Analysis and Applications*, vol. 333, no. 1, pp. 505–522, 2007.

[17] T. S. Hassan, "Oscillation criteria for half-linear dynamic equations on time scales," *Journal of Mathematical Analysis and Applications*, vol. 345, no. 1, pp. 176–185, 2008.

[18] Z. Han, T. Li, S. Sun, and F. Cao, "Oscillation criteria for third order nonlinear delay dynamic equations on time scales," *Annales Polonici Mathematici*, vol. 99, pp. 143–156, 2010.

[19] B. Karpuz, Ö. Öcalan, and S. Öztürk, "Comparison theorems on the oscillation and asymptotic behaviour of higher-order neutral differential equations," *Glasgow Mathematical Journal*, vol. 52, no. 1, pp. 107–114, 2010.

[20] R. M. Mathsen, Q.-R. Wang, and H.-W. Wu, "Oscillation for neutral dynamic functional equations on time scales," *Journal of Difference Equations and Applications*, vol. 10, no. 7, pp. 651–659, 2004.

[21] Z.-Q. Zhu and Q.-R. Wang, "Existence of nonoscillatory solutions to neutral dynamic equations on time scales," *Journal of Mathematical Analysis and Applications*, vol. 335, no. 2, pp. 751–762, 2007.

[22] R. P. Agarwal, D. O'Regan, and S. H. Saker, "Oscillation criteria for second-order nonlinear neutral delay dynamic equations," *Journal of Mathematical Analysis and Applications*, vol. 300, no. 1, pp. 203–217, 2004.

[23] Y. Şahíner, "Oscillation of second-order neutral delay and mixed-type dynamic equations on time scales," *Advances in Difference Equations*, vol. 2006, Article ID 65626, 9 pages, 2006.

[24] S. H. Saker, R. P. Agarwal, and D. O'Regan, "Oscillation results for second-order nonlinear neutral delay dynamic equations on time scales," *Applicable Analysis*, vol. 86, no. 1, pp. 1–17, 2007.

[25] S. H. Saker, "Oscillation of second-order nonlinear neutral delay dynamic equations on time scales," *Journal of Computational and Applied Mathematics*, vol. 187, no. 2, pp. 123–141, 2006.

[26] S. H. Saker and D. O'Regan, "New oscillation criteria for second-order neutral functional dynamic equations via the generalized Riccati substitution," *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, no. 1, pp. 423–434, 2010.

[27] H.-W. Wu, R.-K. Zhuang, and R. M. Mathsen, "Oscillation criteria for second-order nonlinear neutral variable delay dynamic equations," *Applied Mathematics and Computation*, vol. 178, no. 2, pp. 321–331, 2006.

[28] S.-Y. Zhang and Q.-R. Wang, "Oscillation of second-order nonlinear neutral dynamic equations on time scales," *Applied Mathematics and Computation*, vol. 216, no. 10, pp. 2837–2848, 2010.

[29] E. Akin-Bohner and J. Hoffacker, "Oscillation properties of an Emden-Fowler type equation on discrete time scales," *Journal of Difference Equations and Applications*, vol. 9, no. 6, pp. 603–612, 2003.

[30] E. Akin-Bohner, M. Bohner, and S. H. Saker, "Oscillation criteria for a certain class of second order Emden-Fowler dynamic equations," *Electronic Transactions on Numerical Analysis*, vol. 27, pp. 1–12, 2007.

[31] S. R. Grace, R. P. Agarwal, M. Bohner, and D. O'Regan, "Oscillation of second-order strongly superlinear and strongly sublinear dynamic equations," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 8, pp. 3463–3471, 2009.

to Contents, page 1

[32] Z. Han, S. Sun, and B. Shi, “Oscillation criteria for a class of second-order Emden-Fowler delay dynamic equations on time scales,” *Journal of Mathematical Analysis and Applications*, vol. 334, no. 2, pp. 847–858, 2007.
[33] S. Sun, Z. Han, P. Zhao, and C. Zhang, “Oscillation for a class of second-order Emden-Fowler delay dynamic equations on time scales,” *Advances in Difference Equations*, vol. 2010, Article ID 642356, 15 pages, 2010.
[34] S. H. Saker, “Oscillation of second-order neutral delay dynamic equations of Emden-Fowler type,” *Dynamic Systems and Applications*, vol. 15, no. 3-4, pp. 629–644, 2006.
[35] H. A. Agwo, “Oscillation of nonlinear second order neutral delay dynamic equations on time scales,” *Bulletin of the Korean Mathematical Society*, vol. 45, no. 2, pp. 299–312, 2008.
[36] Z. Han, T. Li, S. Sun, and C. Zhang, “Oscillation behavior of third-order neutral Emden-Fowler delay dynamic equations on time scales,” *Advances in Difference Equations*, vol. 2010, Article ID 586312, 23 pages, 2010.
[37] S. H. Abdallah, “Oscillatory and non-oscillatory behaviour of second-order neutral delay differential equations,” *Applied Mathematics and Computation*, vol. 135, no. 2-3, pp. 333–344, 2003.
[38] X. Lin, “Oscillation of second-order nonlinear neutral differential equations,” *Journal of Mathematical Analysis and Applications*, vol. 309, no. 2, pp. 442–452, 2005.
[39] J. S. W. Wong, “On the generalized Emden-Fowler equation,” *SIAM Review*, vol. 17, pp. 339–360, 1975.
[40] W.-T. Li and S. H. Saker, “Oscillation of second-order sublinear neutral delay difference equations,” *Applied Mathematics and Computation*, vol. 146, no. 2-3, pp. 543–551, 2003.
[41] M. K. Yildiz and Ö. Öcalan, “Oscillation results for higher order nonlinear neutral delay difference equations,” *Applied Mathematics Letters*, vol. 20, no. 3, pp. 243–247, 2007.
[42] J. Cheng, “Kamenev-type oscillation criteria for delay difference equations,” *Acta Mathematica Scientia*, vol. 27, no. 3, pp. 574–580, 2007.
[43] H.-F. Huo and W.-T. Li, “Oscillation of the Emden-Fowler difference systems,” *Journal of Mathematical Analysis and Applications*, vol. 256, no. 2, pp. 478–485, 2001.
[44] W.-T. Li and S. S. Cheng, “Asymptotically linear solutions of a discrete Emden-Fowler equation,” *Far East Journal of Mathematical Sciences*, vol. 6, no. 4, pp. 521–542, 1998.
[45] S. H. Saker, “Kamenev-type oscillation criteria for forced Emden-Fowler superlinear difference equations,” *Electronic Journal of Differential Equations*, vol. 2006, no. 68, pp. 1–9, 2002.
[46] J. S. W. Wong, “Necessary and sufficient conditions for oscillation of second order neutral differential equations [Journal of Mathematical Analysis and Applications, vol. 252, no.1, pp. 342–352, 2002],” *Journal of Mathematical Analysis and Applications*, vol. 252, no. 1, pp. 342–352, 2002.
[47] J. S. W. Wong, “Corrigendum to: “Necessary and sufficient conditions for oscillation of second order neutral differential equations”,” *Journal of Mathematical Analysis and Applications*, vol. 332, no. 1, pp. 751–752, 2007.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 367541, 17 pages
doi:10.1155/2011/367541

Research Article

Oscillation Criteria for Second-Order Superlinear Neutral Differential Equations

Tongxing Li,[1,2] Zhenlai Han,[1,2] Chenghui Zhang,[2] and Hua Li[1]

[1] *School of Science, University of Jinan, Jinan, Shandong 250022, China*
[2] *School of Control Science and Engineering, Shandong University, Jinan, Shandong 250061, China*

Correspondence should be addressed to Zhenlai Han, hanzhenlai@163.com

Received 5 September 2010; Accepted 20 January 2011

Academic Editor: Josef Diblík

Some oscillation criteria are established for the second-order superlinear neutral differential equations $(r(t)|z'(t)|^{\alpha-1}z'(t))' + f(t, x(\sigma(t))) = 0$, $t \geq t_0$, where $z(t) = x(t) + p(t)x(\tau(t))$, $\tau(t) \geq t$, $\sigma(t) \geq t$, $p \in C([t_0, \infty), [0, p_0])$, and $\alpha \geq 1$. Our results are based on the cases $\int_{t_0}^{\infty} 1/r^{1/\alpha}(t)\mathrm{d}t = \infty$ or $\int_{t_0}^{\infty} 1/r^{1/\alpha}(t)\mathrm{d}t < \infty$. Two examples are also provided to illustrate these results.

1. Introduction

This paper is concerned with the oscillatory behavior of the second-order superlinear differential equation

$$\left(r(t)\left|z'(t)\right|^{\alpha-1}z'(t)\right)' + f(t, x(\sigma(t))) = 0, \quad t \geq t_0, \tag{1.1}$$

where $\alpha \geq 1$ is a constant, $z(t) = x(t) + p(t)x(\tau(t))$.

Throughout this paper, we will assume the following hypotheses:

(A_1) $r \in C^1([t_0, \infty), \mathbb{R})$, $r(t) > 0$ for $t \geq t_0$;

(A_2) $p \in C([t_0, \infty), [0, p_0])$, where p_0 is a constant;

(A_3) $\tau \in C^1([t_0, \infty), \mathbb{R})$, $\tau'(t) \geq \tau_0 > 0$, $\tau(t) \geq t$;

(A_4) $\sigma \in C([t_0, \infty), \mathbb{R})$, $\sigma(t) \geq t, \tau \circ \sigma = \sigma \circ \tau$;

to Contents, page 1

(A_5) $f(t,u) \in C([t_0,\infty)\times\mathbb{R},\mathbb{R})$, and there exists a function $q \in C([t_0,\infty),[0,\infty))$ such that

$$f(t,u)\operatorname{sign} u \geq q(t)|u|^{\alpha}, \quad \text{for } u \neq 0,\ t \geq t_0. \tag{1.2}$$

By a solution of (1.1), we mean a function $x \in C([T_x,\infty),\mathbb{R})$ for some $T_x \geq t_0$ which has the property that $r(t)|z'(t)|^{\alpha-1}z'(t) \in C^1([T_x,\infty),\mathbb{R})$ and satisfies (1.1) on $[T_x,\infty)$. We consider only those solutions x which satisfy $\sup\{|x(t)| : t \geq T\} > 0$ for all $T \geq T_x$. As is customary, a solution of (1.1) is called oscillatory if it has arbitrarily large zeros on $[t_0,\infty)$, otherwise, it is called nonoscillatory. Equation (1.1) is said to be oscillatory if all its solutions are oscillatory.

We note that neutral differential equations find numerous applications in electric networks. For instance, they are frequently used for the study of distributed networks containing lossless transmission lines which rise in high-speed computers where the lossless transmission lines are used to interconnect switching circuits; see [1].

In the last few years, there are many studies that have been made on the oscillation and asymptotic behavior of solutions of discrete and continuous equations; see, for example, [2–28]. Agarwal et al. [5], Chern et al. [6], Džurina and Stavroulakis [7], Kusano and Yoshida [8], Kusano and Naito [9], Mirzov [10], and Sun and Meng [11] observed some similar properties between

$$\left(r(t)\left|x'(t)\right|^{\alpha-1}x'(t)\right)' + q(t)|x(\sigma(t))|^{\alpha-1}x(\sigma(t)) = 0 \tag{1.3}$$

and the corresponding linear equations

$$\left(r(t)x'(t)\right)' + q(t)x(t) = 0. \tag{1.4}$$

Baculíková [12] established some new oscillation results for (1.3) when $\alpha = 1$. In 1989, Philos [13] obtained some Philos-type oscillation criteria for (1.4).

Recently, many results have been obtained on oscillation and nonoscillation of neutral differential equations, and we refer the reader to [14–35] and the references cited therein. Liu and Bai [16], Xu and Meng [17, 18], Dong [19], Baculíková and Lacková [20], and Jiang and Li [21] established some oscillation criteria for (1.3) with neutral term under the assumptions $\tau(t) \leq t$, $\sigma(t) \leq t$,

$$R(t) = \int_{t_0}^{t} \frac{1}{r^{1/\alpha}(s)}\mathrm{d}s \longrightarrow \infty \text{ as } t \longrightarrow \infty, \tag{1.5}$$

$$\int_{t_0}^{\infty} \frac{1}{r^{1/\alpha}(t)}\mathrm{d}t < \infty. \tag{1.6}$$

Saker and O'Regan [24] studied the oscillatory behavior of (1.1) when $0 \leq p(t) < 1$, $\tau(t) \leq t$ and $\sigma(t) > t$.

Han et al. [26] examined the oscillation of second-order nonlinear neutral differential equation

$$\left(r(t)\left[x(t)+p(t)x(\tau(t))\right]'\right)' + q(t)f(x(\sigma(t))) = 0, \quad t \geq t_0, \tag{1.7}$$

to Contents, page 1

where $\tau(t) \leq t$, $\sigma(t) \leq t$, $\tau'(t) = \tau_0 > 0$, $0 \leq p(t) \leq p_0 < \infty$, and the authors obtained some oscillation criteria for (1.7).

However, there are few results regarding the oscillatory problem of (1.1) when $\tau(t) \geq t$ and $\sigma(t) \geq t$. Our aim in this paper is to establish some oscillation criteria for (1.1) under the case when $\tau(t) \geq t$ and $\sigma(t) \geq t$.

The paper is organized as follows. In Section 2, we will establish an inequality to prove our results. In Section 3, some oscillation criteria are obtained for (1.1). In Section 4, we give two examples to show the importance of the main results.

All functional inequalities considered in this paper are assumed to hold eventually, that is, they are satisfied for all t large enough.

2. Lemma

In this section, we give the following lemma, which we will use in the proofs of our main results.

Lemma 2.1. *Assume that $\alpha \geq 1$, $a, b \in \mathbb{R}$. If $a \geq 0$, $b \geq 0$, then*

$$a^\alpha + b^\alpha \geq \frac{1}{2^{\alpha-1}}(a+b)^\alpha \tag{2.1}$$

holds.

Proof. (i) Suppose that $a = 0$ or $b = 0$. Obviously, we have (2.1). (ii) Suppose that $a > 0$, $b > 0$. Define the function g by $g(u) = u^\alpha$, $u \in (0, \infty)$. Then $g''(u) = \alpha(\alpha - 1)u^{\alpha-2} \geq 0$ for $u > 0$. Thus, g is a convex function. By the definition of convex function, for $\lambda = 1/2$, $a, b \in (0, \infty)$, we have

$$g\left(\frac{a+b}{2}\right) \leq \frac{g(a) + g(b)}{2}, \tag{2.2}$$

that is,

$$a^\alpha + b^\alpha \geq \frac{1}{2^{\alpha-1}}(a+b)^\alpha. \tag{2.3}$$

This completes the proof. □

3. Main Results

In this section, we will establish some oscillation criteria for (1.1).

First, we establish two comparison theorems which enable us to reduce the problem of the oscillation of (1.1) to the research of the first-order differential inequalities.

to Contents, page 1

Theorem 3.1. *Suppose that* (1.5) *holds. If the first-order neutral differential inequality*

$$\left[u(t)+\frac{(p_0)^{\alpha}}{\tau_0}u(\tau(t))\right]'+\frac{1}{2^{\alpha-1}}Q(t)(R(\sigma(t))-R(t_1))^{\alpha}u(\sigma(t))\leq 0 \tag{3.1}$$

has no positive solution for all sufficiently large t_1, *where* $Q(t) = \min\{q(t), q(\tau(t))\}$, *then every solution of* (1.1) *oscillates.*

Proof. Let x be a nonoscillatory solution of (1.1). Without loss of generality, we assume that there exists $t_1 \geq t_0$ such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\sigma(t)) > 0$ for all $t \geq t_1$. Then $z(t) > 0$ for $t \geq t_1$. In view of (1.1), we obtain

$$\left(r(t)\left|z'(t)\right|^{\alpha-1}z'(t)\right)'\leq -q(t)x^{\alpha}(\sigma(t))\leq 0,\quad t\geq t_1. \tag{3.2}$$

Thus, $r(t)|z'(t)|^{\alpha-1}z'(t)$ is decreasing function. Now we have two possible cases for $z'(t)$: (i) $z'(t) < 0$ eventually, (ii) $z'(t) > 0$ eventually.

Suppose that $z'(t) < 0$ for $t \geq t_2 \geq t_1$. Then, from (3.2), we get

$$r(t)\left|z'(t)\right|^{\alpha-1}z'(t)\leq r(t_2)\left|z'(t_2)\right|^{\alpha-1}z'(t_2),\quad t\geq t_2, \tag{3.3}$$

which implies that

$$z(t)\leq z(t_2)-r^{1/\alpha}(t_2)\left|z'(t_2)\right|\int_{t_2}^{t}r^{-1/\alpha}(s)\mathrm{d}s. \tag{3.4}$$

Letting $t \to \infty$, by (1.5), we find $z(t) \to -\infty$, which is a contradiction. Thus

$$z'(t)>0 \tag{3.5}$$

for $t \geq t_2$.

By applying (1.1), for all sufficiently large t, we obtain

$$\left(r(t)\left(z'(t)\right)^{\alpha}\right)'+q(t)x^{\alpha}(\sigma(t))+(p_0)^{\alpha}q(\tau(t))x^{\alpha}(\sigma(\tau(t)))+\frac{(p_0)^{\alpha}}{\tau'(t)}\left(r(\tau(t))\left(z'(\tau(t))\right)^{\alpha}\right)'\leq 0. \tag{3.6}$$

Using inequality (2.1), (3.2), (3.5), $\tau \circ \sigma = \sigma \circ \tau$, and the definition of z, we conclude that

$$\left(r(t)\left(z'(t)\right)^{\alpha}\right)'+\frac{(p_0)^{\alpha}}{\tau_0}r(\tau(t))\left(z'(\tau(t))^{\alpha}\right)'+\frac{1}{2^{\alpha-1}}Q(t)z^{\alpha}(\sigma(t))\leq 0. \tag{3.7}$$

to Contents, page 1

It follows from (3.2) and (3.5) that $u(t) = r(t)(z'(t))^{\alpha} > 0$ is decreasing and then

$$z(t) \geq \int_{t_2}^{t} \frac{\left(r(s)(z'(s))^{\alpha}\right)^{1/\alpha}}{r^{1/\alpha}(s)} ds \geq u^{1/\alpha}(t) \int_{t_2}^{t} \frac{1}{r^{1/\alpha}(s)} ds = u^{1/\alpha}(t)(R(t) - R(t_2)). \tag{3.8}$$

Thus, from (3.7) and the above inequality, we find

$$\left[u(t) + \frac{(p_0)^{\alpha}}{\tau_0} u(\tau(t))\right]' + \frac{1}{2^{\alpha-1}} Q(t)(R(\sigma(t)) - R(t_2))^{\alpha} u(\sigma(t)) \leq 0. \tag{3.9}$$

That is, inequality (3.1) has a positive solution u; this is a contradiction. The proof is complete. □

Theorem 3.2. *Suppose that* (1.5) *holds. If the first-order differential inequality*

$$\eta'(t) + \frac{\tau_0}{2^{\alpha-1}\left(\tau_0 + (p_0)^{\alpha}\right)} Q(t)(R(\sigma(t)) \ - R(t_1))^{\alpha} \eta(\sigma(t)) \leq 0 \tag{3.10}$$

has no positive solution for all sufficiently large t_1, where Q is defined as in Theorem 3.1, then every solution of (1.1) *oscillates.*

Proof. Let x be a nonoscillatory solution of (1.1). Without loss of generality, we assume that there exists $t_1 \geq t_0$ such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\sigma(t)) > 0$ for all $t \geq t_1$. Then $z(t) > 0$ for $t \geq t_1$. Proceeding as in the proof of Theorem 3.1, we obtain that $u(t) = r(t)(z'(t))^{\alpha}$ is decreasing, and it satisfies inequality (3.1). Set $\eta(t) = u(t) + (p_0)^{\alpha} u(\tau(t))/\tau_0$. From $\tau(t) \geq t$, we get

$$\eta(t) = u(t) + \frac{(p_0)^{\alpha}}{\tau_0} u(\tau(t)) \leq \left(1 + \frac{(p_0)^{\alpha}}{\tau_0}\right) u(t). \tag{3.11}$$

In view of the above inequality and (3.1), we see that

$$\eta'(t) + \frac{\tau_0}{2^{\alpha-1}\left(\tau_0 + (p_0)^{\alpha}\right)} Q(t)(R(\sigma(t)) - R(t_1))^{\alpha} \eta(\sigma(t)) \leq 0. \tag{3.12}$$

That is, inequality (3.10) has a positive solution η; this is a contradiction. The proof is complete. □

Next, using Riccati transformation technique, we obtain the following results.

Theorem 3.3. *Suppose that* (1.5) *holds. Moreover, assume that there exists $\rho \in C^1([t_0, \infty), (0, \infty))$ such that*

$$\limsup_{t \to \infty} \int_{t_0}^{t} \left[\frac{\rho(s) Q(s)}{2^{\alpha-1}} - \frac{1}{(\alpha+1)^{\alpha+1}} \left(1 + \frac{(p_0)^{\alpha}}{\tau_0}\right) \frac{r(s)\left(\rho'_+(s)\right)^{\alpha+1}}{(\rho(s))^{\alpha}}\right] ds = \infty \tag{3.13}$$

to Contents, page 1

holds, where Q *is defined as in Theorem 3.1,* $\rho'_+(t) = \max\{0, \rho'(t)\}$*. Then every solution of* (1.1) *oscillates.*

Proof. Let x be a nonoscillatory solution of (1.1). Without loss of generality, we assume that there exists $t_1 \ge t_0$ such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\sigma(t)) > 0$ for all $t \ge t_1$. Then $z(t) > 0$ for $t \ge t_1$. Proceeding as in the proof of Theorem 3.1, we obtain (3.2)–(3.7).

Define a Riccati substitution

$$\omega(t) = \rho(t)\frac{r(t)(z'(t))^{\alpha}}{(z(t))^{\alpha}}, \quad t \ge t_2. \tag{3.14}$$

Thus $\omega(t) > 0$ for $t \ge t_2$. Differentiating (3.14) we find that

$$\omega'(t) = \rho'(t)\frac{r(t)(z'(t))^{\alpha}}{(z(t))^{\alpha}} + \rho(t)\frac{\left(r(t)(z'(t))^{\alpha}\right)'}{(z(t))^{\alpha}} - \alpha\rho(t)\frac{r(t)(z'(t))^{\alpha}z^{\alpha-1}(t)z'(t)}{(z(t))^{2\alpha}}. \tag{3.15}$$

Hence, by (3.14) and (3.15), we see that

$$\omega'(t) = \frac{\rho'(t)}{\rho(t)}\omega(t) + \rho(t)\frac{\left(r(t)(z'(t))^{\alpha}\right)'}{(z(t))^{\alpha}} - \frac{\alpha}{\rho^{1/\alpha}(t)r^{1/\alpha}(t)}\omega^{(\alpha+1)/\alpha}(t). \tag{3.16}$$

Similarly, we introduce another Riccati substitution

$$\upsilon(t) = \rho(t)\frac{r(\tau(t))(z'(\tau(t)))^{\alpha}}{(z(t))^{\alpha}}, \quad t \ge t_2. \tag{3.17}$$

Then $\upsilon(t) > 0$ for $t \ge t_2$. From (3.2), (3.5), and $\tau(t) \ge t$, we have

$$z'(t) \ge \left(\frac{r(\tau(t))}{r(t)}\right)^{1/\alpha} z'(\tau(t)). \tag{3.18}$$

Differentiating (3.17), we find

$$\upsilon'(t) = \rho'(t)\frac{r(\tau(t))(z'(\tau(t)))^{\alpha}}{(z(t))^{\alpha}} + \rho(t)\frac{\left(r(\tau(t))(z'(\tau(t)))^{\alpha}\right)'}{(z(t))^{\alpha}} - \alpha\rho(t)\frac{r(\tau(t))(z'(\tau(t)))^{\alpha}z^{\alpha-1}(t)z'(t)}{(z(t))^{2\alpha}}. \tag{3.19}$$

Therefore, by (3.17), (3.18), and (3.19), we see that

$$\upsilon'(t) \le \frac{\rho'(t)}{\rho(t)}\upsilon(t) + \rho(t)\frac{\left(r(\tau(t))(z'(\tau(t)))^{\alpha}\right)'}{(z(t))^{\alpha}} - \frac{\alpha}{\rho^{1/\alpha}(t)r^{1/\alpha}(t)}\upsilon^{(\alpha+1)/\alpha}(t). \tag{3.20}$$

to Contents, page 1

Thus, from (3.16) and (3.20), we have

$$\begin{aligned}\omega'(t)+\frac{(p_0)^{\alpha}}{\tau_0}\upsilon'(t)\le{}&\rho(t)\frac{\left(r(t)(z'(t))^{\alpha}\right)'+\left((p_0)^{\alpha}/\tau_0\right)\left(r(\tau(t))(z'(\tau(t)))^{\alpha}\right)'}{(z(t))^{\alpha}}\\&+\frac{\rho'(t)}{\rho(t)}\omega(t)-\frac{\alpha}{\rho^{1/\alpha}(t)r^{1/\alpha}(t)}\omega^{(\alpha+1)/\alpha}(t)+\frac{(p_0)^{\alpha}}{\tau_0}\frac{\rho'(t)}{\rho(t)}\upsilon(t)\\&-\frac{(p_0)^{\alpha}}{\tau_0}\frac{\alpha}{\rho^{1/\alpha}(t)r^{1/\alpha}(t)}\upsilon^{(\alpha+1)/\alpha}(t).\end{aligned}\tag{3.21}$$

It follows from (3.5), (3.7), and $\sigma(t)\ge t$ that

$$\begin{aligned}\omega'(t)+\frac{(p_0)^{\alpha}}{\tau_0}\upsilon'(t)\le{}&-\frac{1}{2^{\alpha-1}}\rho(t)Q(t)+\frac{\rho'_+(t)}{\rho(t)}\omega(t)-\frac{\alpha}{\rho^{1/\alpha}(t)r^{1/\alpha}(t)}\omega^{(\alpha+1)/\alpha}(t)\\&+\frac{(p_0)^{\alpha}}{\tau_0}\frac{\rho'_+(t)}{\rho(t)}\upsilon(t)-\frac{(p_0)^{\alpha}}{\tau_0}\frac{\alpha}{\rho^{1/\alpha}(t)r^{1/\alpha}(t)}\upsilon^{(\alpha+1)/\alpha}(t).\end{aligned}\tag{3.22}$$

Integrating the above inequality from t_2 to t, we obtain

$$\begin{aligned}&\omega(t)-\omega(t_2)+\frac{(p_0)^{\alpha}}{\tau_0}\upsilon(t)-\frac{(p_0)^{\alpha}}{\tau_0}\upsilon(t_2)\\&\quad\le-\int_{t_2}^{t}\frac{1}{2^{\alpha-1}}\rho(s)Q(s)\mathrm{d}s+\int_{t_2}^{t}\left[\frac{\rho'_+(s)}{\rho(s)}\omega(s)-\frac{\alpha}{\rho^{1/\alpha}(s)r^{1/\alpha}(s)}\omega^{(\alpha+1)/\alpha}(s)\right]\mathrm{d}s\\&\qquad+\int_{t_2}^{t}\frac{(p_0)^{\alpha}}{\tau_0}\left[\frac{\rho'_+(s)}{\rho(s)}\upsilon(s)-\frac{\alpha}{\rho^{1/\alpha}(s)r^{1/\alpha}(s)}\upsilon^{(\alpha+1)/\alpha}(s)\right]\mathrm{d}s.\end{aligned}\tag{3.23}$$

Define

$$A:=\left[\frac{\alpha}{\rho^{1/\alpha}(t)r^{1/\alpha}(t)}\right]^{\alpha/(\alpha+1)}\omega(t),\qquad B:=\left[\frac{\rho'_+(t)}{\rho(t)}\frac{\alpha}{\alpha+1}\left[\frac{\alpha}{\rho^{1/\alpha}(t)r^{1/\alpha}(t)}\right]^{-\alpha/(\alpha+1)}\right]^{\alpha}.\tag{3.24}$$

Using inequality

$$\frac{\alpha+1}{\alpha}AB^{1/\alpha}-A^{(\alpha+1)/\alpha}\le\frac{1}{\alpha}B^{(\alpha+1)/\alpha},\quad\text{for } A\ge 0,\ B\ge 0\ \text{ are constants,}\tag{3.25}$$

we have

$$\frac{\rho'_+(t)}{\rho(t)}\omega(t)-\frac{\alpha}{\rho^{1/\alpha}(t)r^{1/\alpha}(t)}\omega^{(\alpha+1)/\alpha}(t)\le\frac{1}{(\alpha+1)^{\alpha+1}}\frac{r(t)\left(\rho'_+(t)\right)^{\alpha+1}}{\rho(t)^{\alpha}}.\tag{3.26}$$

to Contents, page 1

Similarly, we obtain

$$\frac{\rho'_+(t)}{\rho(t)}v(t) - \frac{\alpha}{\rho^{1/\alpha}(t)r^{1/\alpha}(t)}v^{(\alpha+1)/\alpha}(t) \leq \frac{1}{(\alpha+1)^{\alpha+1}}\frac{r(t)\left(\rho'_+(t)\right)^{\alpha+1}}{\rho(t)^{\alpha}}. \tag{3.27}$$

Thus, from (3.23), we get

$$\begin{aligned} &\omega(t) - \omega(t_2) + \frac{(p_0)^{\alpha}}{\tau_0}v(t) - \frac{(p_0)^{\alpha}}{\tau_0}v(t_2) \\ &\quad\leq -\int_{t_2}^{t}\left[\frac{\rho(s)Q(s)}{2^{\alpha-1}} - \frac{1}{(\alpha+1)^{\alpha+1}}\left(1+\frac{(p_0)^{\alpha}}{\tau_0}\right)\frac{r(s)\left(\rho'_+(s)\right)^{\alpha+1}}{\rho(s)^{\alpha}}\right]\mathrm{d}s, \end{aligned} \tag{3.28}$$

which contradicts (3.13). This completes the proof. □

As an immediate consequence of Theorem 3.3 we get the following.

Corollary 3.4. *Let assumption* (3.13) *in Theorem 3.3 be replaced by*

$$\begin{gathered} \limsup_{t\to\infty}\int_{t_0}^{t}\rho(s)Q(s)\mathrm{d}s = \infty, \\ \limsup_{t\to\infty}\int_{t_0}^{t}\frac{r(s)\left(\rho'_+(s)\right)^{\alpha+1}}{\left(\rho(s)\right)^{\alpha}}\mathrm{d}s < \infty. \end{gathered} \tag{3.29}$$

Then every solution of (1.1) *oscillates.*

From Theorem 3.3 by choosing the function ρ, appropriately, we can obtain different sufficient conditions for oscillation of (1.1), and if we define a function ρ by $\rho(t) = 1$, and $\rho(t) = t$, we have the following oscillation results.

Corollary 3.5. *Suppose that* (1.5) *holds. If*

$$\limsup_{t\to\infty}\int_{t_0}^{t}Q(s)\mathrm{d}s = \infty, \tag{3.30}$$

where Q is defined as in Theorem 3.1, then every solution of (1.1) *oscillates.*

Corollary 3.6. *Suppose that* (1.5) *holds. If*

$$\limsup_{t\to\infty}\int_{t_0}^{t}\left[\frac{sQ(s)}{2^{\alpha-1}} - \frac{1}{(\alpha+1)^{\alpha+1}}\left(1+\frac{(p_0)^{\alpha}}{\tau_0}\right)\frac{r(s)}{s^{\alpha}}\right]\mathrm{d}s = \infty, \tag{3.31}$$

where Q is defined as in Theorem 3.1, then every solution of (1.1) *oscillates.*

to Contents, page 1

In the following theorem, we present a Philos-type oscillation criterion for (1.1). First, we introduce a class of functions $\mathfrak{R}$. Let

$$\mathbb{D}_0 = \{(t,s) : t > s \geq t_0\}, \qquad \mathbb{D} = \{(t,s) : t \geq s \geq t_0\}. \tag{3.32}$$

The function $H \in C(\mathbb{D}, \mathbb{R})$ is said to belong to the class $\mathfrak{R}$ (defined by $H \in \mathfrak{R}$, for short) if

(i) $H(t,t) = 0$, for $t \geq t_0$, $H(t,s) > 0$, for $(t,s) \in \mathbb{D}_0$;

(ii) H has a continuous and nonpositive partial derivative $\partial H(t,s)/\partial s$ on D_0 with respect to s.

We assume that $\varsigma(t)$ and $\rho(t)$ for $t \geq t_0$ are given continuous functions such that $\rho(t) > 0$ and differentiable and define

$$\theta(t) = \frac{\rho'(t)}{\rho(t)} + (\alpha+1)(\varsigma(t))^{1/\alpha}, \psi(t) = \rho(t)\left\{[r(t)\varsigma(t)]' - r(t)(\varsigma(t))^{(1+\alpha)/\alpha}\right\},$$
$$\phi(t,s) = \frac{r(s)\rho(s)}{(\alpha+1)^{\alpha+1}}\left(\theta(s) + \frac{\partial H(t,s)/\partial s}{H(t,s)}\right)^{\alpha+1}. \tag{3.33}$$

Now, we give the following result.

Theorem 3.7. *Suppose that* (1.5) *holds and α is a quotient of odd positive integers. Moreover, let $H \in \mathfrak{R}$ be such that*

$$\limsup_{t\to\infty} \frac{1}{H(t,t_0)} \int_{t_0}^{t} H(t,s)\left[\frac{\rho(s)Q(s)}{2^{\alpha-1}} - \left(1 + \frac{(p_0)^\alpha}{\tau_0}\right)(\psi(s) + \phi(t,s))\right] \mathrm{d}s = \infty \tag{3.34}$$

holds, where Q is defined as in Theorem 3.1. Then every solution of (1.1) *oscillates.*

Proof. Let x be a nonoscillatory solution of (1.1). Without loss of generality, we assume that there exists $t_1 \geq t_0$ such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\sigma(t)) > 0$ for all $t \geq t_1$. Then $z(t) > 0$ for $t \geq t_1$. Proceeding as in the proof of Theorem 3.1, we obtain (3.2)–(3.7). Define the Riccati substitution ω by

$$\omega(t) = \rho(t)\left[\frac{r(t)(z'(t))^\alpha}{(z(t))^\alpha} + r(t)\varsigma(t)\right], \quad t \geq t_2 \geq t_1. \tag{3.35}$$

Then, we have

$$\omega'(t) = \rho'(t)\left[\frac{r(t)(z'(t))^\alpha}{(z(t))^\alpha} + r(t)\varsigma(t)\right] + \rho(t)\left[\frac{r(t)(z'(t))^\alpha}{(z(t))^\alpha} + r(t)\varsigma(t)\right]'$$
$$= \frac{\rho'(t)}{\rho(t)}\omega(t) + \rho(t)[r(t)\varsigma(t)]' + \rho(t)\frac{\left(r(t)(z'(t))^\alpha\right)'}{(z(t))^\alpha} - \alpha\rho(t)\frac{r(t)(z'(t))^{\alpha+1}}{(z(t))^{\alpha+1}}. \tag{3.36}$$

to Contents, page 1

Using (3.35), we get

$$\omega'(t) = \frac{\rho'(t)}{\rho(t)}\omega(t) + \rho(t)[r(t)\varsigma(t)]' + \rho(t)\frac{\left(r(t)(z'(t))^{\alpha}\right)'}{(z(t))^{\alpha}} - \frac{\alpha\rho(t)}{r^{1/\alpha}(t)}\left[\frac{\omega(t)}{\rho(t)} - r(t)\varsigma(t)\right]^{(1+\alpha)/\alpha}. \tag{3.37}$$

Let

$$A = \frac{\omega(t)}{\rho(t)}, \qquad B = r(t)\varsigma(t). \tag{3.38}$$

By applying the inequality (see [21, 24])

$$A^{(1+\alpha)/\alpha} - (A-B)^{1+\alpha/\alpha} \le B^{1/\alpha}\left[\left(1+\frac{1}{\alpha}\right)A - \frac{1}{\alpha}B\right], \quad \text{for } \alpha = \frac{\text{odd}}{\text{odd}} \ge 1, \tag{3.39}$$

we see that

$$\left[\frac{\omega(t)}{\rho(t)} - r(t)\varsigma(t)\right]^{(1+\alpha)/\alpha} \ge \frac{\omega^{(1+\alpha)/\alpha}(t)}{\rho^{(1+\alpha)/\alpha}(t)} + \frac{1}{\alpha}(r(t)\varsigma(t))^{(1+\alpha)/\alpha} - \frac{\alpha+1}{\alpha}\frac{(r(t)\varsigma(t))^{1/\alpha}}{\rho(t)}\omega(t). \tag{3.40}$$

Substituting (3.40) into (3.37), we have

$$\begin{aligned}\omega'(t) \le{}& \left[\frac{\rho'(t)}{\rho(t)} + (\alpha+1)(\varsigma(t))^{1/\alpha}\right]\omega(t) + \rho(t)\left\{[r(t)\varsigma(t)]' - r(t)(\varsigma(t))^{(1+\alpha)/\alpha}\right\}\\ &+ \rho(t)\frac{\left(r(t)(z'(t))^{\alpha}\right)'}{(z(t))^{\alpha}} - \frac{\alpha}{r^{1/\alpha}(t)\rho^{1/\alpha}(t)}\omega^{(1+\alpha)/\alpha}(t).\end{aligned} \tag{3.41}$$

That is,

$$\omega'(t) \le \theta(t)\omega(t) + \psi(t) + \rho(t)\frac{\left(r(t)(z'(t))^{\alpha}\right)'}{(z(t))^{\alpha}} - \frac{\alpha}{r^{1/\alpha}(t)\rho^{1/\alpha}(t)}\omega^{(1+\alpha)/\alpha}(t). \tag{3.42}$$

Next, define another Riccati transformation u by

$$u(t) = \rho(t)\left[\frac{r(\tau(t))(z'(\tau(t)))^{\alpha}}{(z(t))^{\alpha}} + r(t)\varsigma(t)\right], \quad t \ge t_2 \ge t_1. \tag{3.43}$$

to Contents, page 1

Then, we have

$$u'(t) = \rho'(t)\left[\frac{r(\tau(t))(z'(\tau(t)))^{\alpha}}{(z(t))^{\alpha}} + r(t)\varsigma(t)\right] + \rho(t)\left[\frac{r(\tau(t))(z'(\tau(t)))^{\alpha}}{(z(t))^{\alpha}} + r(t)\varsigma(t)\right]'$$
$$= \frac{\rho'(t)}{\rho(t)}u(t) + \rho(t)[r(t)\varsigma(t)]' + \rho(t)\frac{\left(r(\tau(t))(z'(\tau(t)))^{\alpha}\right)'}{(z(t))^{\alpha}} - \alpha\rho(t)\frac{r(\tau(t))(z'(\tau(t)))^{\alpha}z'(t)}{(z(t))^{\alpha+1}}. \tag{3.44}$$

From (3.2), (3.5), and $\tau(t) \geq t$, we have that (3.18) holds. Hence, we obtain

$$u'(t) \leq \frac{\rho'(t)}{\rho(t)}u(t) + \rho(t)[r(t)\varsigma(t)]' + \rho(t)\frac{\left(r(\tau(t))(z'(\tau(t)))^{\alpha}\right)}{(z(t))^{\alpha}} - \alpha\rho(t)\frac{\left(r(\tau(t))(z'(\tau(t)))^{\alpha}\right)^{(1+\alpha)/\alpha}}{r^{1/\alpha}(t)(z(t))^{\alpha+1}}. \tag{3.45}$$

Using (3.43), we get

$$u'(t) \leq \frac{\rho'(t)}{\rho(t)}u(t) + \rho(t)[r(t)\varsigma(t)]' + \rho(t)\frac{\left(r(\tau(t))(z'(\tau(t)))^{\alpha}\right)'}{(z(t))^{\alpha}} - \frac{\alpha\rho(t)}{r^{1/\alpha}(t)}\left[\frac{u(t)}{\rho(t)} - r(t)\varsigma(t)\right]^{(1+\alpha)/\alpha}. \tag{3.46}$$

Let

$$A = \frac{u(t)}{\rho(t)}, \qquad B = r(t)\varsigma(t). \tag{3.47}$$

By applying the inequality (3.39), we see that

$$\left[\frac{u(t)}{\rho(t)} - r(t)\varsigma(t)\right]^{(1+\alpha)/\alpha} \geq \frac{u^{(1+\alpha)/\alpha}(t)}{\rho^{(1+\alpha)/\alpha}(t)} + \frac{1}{\alpha}(r(t)\varsigma(t))^{(1+\alpha)/\alpha} - \frac{\alpha+1}{\alpha}\frac{(r(t)\varsigma(t))^{1/\alpha}}{\rho(t)}u(t). \tag{3.48}$$

Substituting (3.48) into (3.46), we have

$$\begin{aligned} u'(t) \leq{}& \left[\frac{\rho'(t)}{\rho(t)} + (\alpha+1)(\varsigma(t))^{1/\alpha}\right]u(t) + \rho(t)\left\{[r(t)\varsigma(t)]' - r(t)(\varsigma(t))^{(1+\alpha)/\alpha}\right\} \\ &+ \rho(t)\frac{\left(r(\tau(t))(z'(\tau(t)))^{\alpha}\right)'}{(z(t))^{\alpha}} - \frac{\alpha}{r^{1/\alpha}(t)\rho^{1/\alpha}(t)}u^{(1+\alpha)/\alpha}(t). \end{aligned} \tag{3.49}$$

That is,

$$u'(t) \leq \theta(t)u(t) + \psi(t) + \rho(t)\frac{\left(r(\tau(t))(z'(\tau(t)))^{\alpha}\right)'}{(z(t))^{\alpha}} - \frac{\alpha}{r^{1/\alpha}(t)\rho^{1/\alpha}(t)}u^{(1+\alpha)/\alpha}(t). \tag{3.50}$$

to Contents, page 1

By (3.42) and (3.50), we find

$$\begin{aligned}
\omega'(t) + \frac{(p_0)^\alpha}{\tau_0}u'(t) \le{}& \left(1 + \frac{(p_0)^\alpha}{\tau_0}\right)\varphi(t) + \rho(t)\frac{\left(r(t)(z'(t))^\alpha\right)' + \left((p_0)^\alpha/\tau_0\right)\left(r(\tau(t))(z'(\tau(t)))^\alpha\right)'}{(z(t))^\alpha} \\
&+ \theta(t)\omega(t) - \frac{\alpha}{r^{1/\alpha}(t)\rho^{1/\alpha}(t)}\omega^{(1+\alpha)/\alpha}(t) + \frac{(p_0)^\alpha}{\tau_0}\theta(t)u(t) \\
&- \frac{(p_0)^\alpha}{\tau_0}\frac{\alpha}{r^{1/\alpha}(t)\rho^{1/\alpha}(t)}u^{(1+\alpha)/\alpha}(t).
\end{aligned} \tag{3.51}$$

In view of the above inequality, (3.5), (3.7), and $\sigma(t) \ge t$, we get

$$\begin{aligned}
\omega'(t) + \frac{(p_0)^\alpha}{\tau_0}u'(t) \le{}& \left(1 + \frac{(p_0)^\alpha}{\tau_0}\right)\varphi(t) - \frac{\rho(t)Q(t)}{2^{\alpha-1}} + \theta(t)\omega(t) - \frac{\alpha}{r^{1/\alpha}(t)\rho^{1/\alpha}(t)}\omega^{(1+\alpha)/\alpha}(t) \\
&+ \frac{(p_0)^\alpha}{\tau_0}\theta(t)u(t) - \frac{(p_0)^\alpha}{\tau_0}\frac{\alpha}{r^{1/\alpha}(t)\rho^{1/\alpha}(t)}u^{(1+\alpha)/\alpha}(t),
\end{aligned} \tag{3.52}$$

which follows that

$$\begin{aligned}
&\int_{t_2}^{t} H(t,s)\left[\frac{\rho(s)Q(s)}{2^{\alpha-1}} - \left(1 + \frac{(p_0)^\alpha}{\tau_0}\right)\varphi(s)\right]\mathrm{d}s \\
&\quad\le -\int_{t_2}^{t} H(t,s)\omega'(s)\mathrm{d}s + \int_{t_2}^{t} H(t,s)\theta(s)\omega(s)\mathrm{d}s \\
&\quad\quad - \int_{t_2}^{t} H(t,s)\frac{\alpha\omega^{(1+\alpha)/\alpha}(s)}{r^{1/\alpha}(s)\rho^{1/\alpha}(s)}\mathrm{d}s - \frac{(p_0)^\alpha}{\tau_0}\int_{t_2}^{t} H(t,s)u'(s)\mathrm{d}s \\
&\quad\quad + \frac{(p_0)^\alpha}{\tau_0}\int_{t_2}^{t} H(t,s)\theta(s)u(s)\mathrm{d}s - \frac{(p_0)^\alpha}{\tau_0}\int_{t_2}^{t} H(t,s)\frac{\alpha u^{(1+\alpha)/\alpha}(s)}{r^{1/\alpha}(s)\rho^{1/\alpha}(s)}\mathrm{d}s.
\end{aligned} \tag{3.53}$$

Using the integration by parts formula and $H(t,t) = 0$, we have

$$\begin{aligned}
\int_{t_2}^{t} H(t,s)\omega'(s)\mathrm{d}s &= -H(t,t_2)\omega(t_2) - \int_{t_2}^{t} \frac{\partial H(t,s)}{\partial s}\omega(s)\mathrm{d}s, \\
\int_{t_2}^{t} H(t,s)u'(s)\mathrm{d}s &= -H(t,t_2)u(t_2) - \int_{t_2}^{t} \frac{\partial H(t,s)}{\partial s}u(s)\mathrm{d}s.
\end{aligned} \tag{3.54}$$

to Contents, page 1

So, by (3.53), we obtain

$$\begin{aligned}
&\int_{t_2}^{t} H(t,s)\left[\frac{\rho(s)Q(s)}{2^{\alpha-1}} - \left(1+\frac{(p_0)^\alpha}{\tau_0}\right)\varphi(s)\right]\mathrm{d}s \\
&\quad \le H(t,t_2)\omega(t_2) + \frac{(p_0)^\alpha}{\tau_0}H(t,t_2)u(t_2) \\
&\qquad + \int_{t_2}^{t} H(t,s)\left[\theta(s)+\frac{\partial H(t,s)/\partial s}{H(t,s)}\right]\omega(s)\mathrm{d}s - \int_{t_2}^{t} H(t,s)\frac{\alpha\omega^{(1+\alpha)/\alpha}(s)}{r^{1/\alpha}(s)\rho^{1/\alpha}(s)}\mathrm{d}s \\
&\qquad + \frac{(p_0)^\alpha}{\tau_0}\int_{t_2}^{t} H(t,s)\left[\theta(s)+\frac{\partial H(t,s)/\partial s}{H(t,s)}\right]u(s)\mathrm{d}s - \frac{(p_0)^\alpha}{\tau_0}\int_{t_2}^{t} H(t,s)\frac{\alpha u^{(1+\alpha)/\alpha}(s)}{r^{1/\alpha}(s)\rho^{1/\alpha}(s)}\mathrm{d}s.
\end{aligned} \tag{3.55}$$

Using the inequality

$$By - Ay^{(\alpha+1)/\alpha} \le \frac{\alpha^\alpha}{(\alpha+1)^{\alpha+1}}\frac{B^{\alpha+1}}{A^\alpha}, \tag{3.56}$$

where

$$A = \frac{\alpha}{r^{1/\alpha}(s)\rho^{1/\alpha}(s)}, \qquad B = \theta(s) + \frac{\partial H(t,s)/\partial s}{H(t,s)}, \tag{3.57}$$

we have

$$\int_{t_2}^{t} H(t,s)\left[\frac{\rho(s)Q(s)}{2^{\alpha-1}} - \left(1+\frac{(p_0)^\alpha}{\tau_0}\right)(\varphi(s)+\phi(t,s))\right]\mathrm{d}s \le H(t,t_2)\omega(t_2) + \frac{(p_0)^\alpha}{\tau_0}H(t,t_2)u(t_2) \tag{3.58}$$

due to (3.55), which yields that

$$\frac{1}{H(t,t_2)}\int_{t_2}^{t} H(t,s)\left[\frac{\rho(s)Q(s)}{2^{\alpha-1}} - \left(1+\frac{(p_0)^\alpha}{\tau_0}\right)(\varphi(s)+\phi(t,s))\right]\mathrm{d}s \le \omega(t_2) + \frac{(p_0)^\alpha}{\tau_0}u(t_2), \tag{3.59}$$

which contradicts (3.34). The proof is complete. □

From Theorem 3.7, we can obtain different oscillation conditions for all solutions of (1.1) with different choices of H; the details are left to the reader.

Theorem 3.8. *Assume that* (1.6) *and* (3.30) *hold. Furthermore, assume that* $0 \le p(t) \le p_1 < 1$. *If*

$$\int_{t_0}^{\infty}\left[\frac{1}{r(s)}\int_{t_0}^{s} q(u)\mathrm{d}u\right]^{1/\alpha}\mathrm{d}s = \infty, \tag{3.60}$$

then every solution x *of* (1.1) *oscillates or* $\lim_{t\to\infty} x(t) = 0$.

Proof. Let x be a nonoscillatory solution of (1.1). Without loss of generality, we assume that there exists $t_1 \geq t_0$ such that $x(t) > 0$, $x(\tau(t)) > 0$, and $x(\sigma(t)) > 0$ for all $t \geq t_1$. Then $z(t) > 0$ for $t \geq t_1$. Proceeding as in the proof of Theorem 3.1, we obtain (3.2). Thus $r(t)|z'(t)|^{\alpha-1}z'(t)$ is decreasing function, and there exists a $t_2 \geq t_1$ such that $z'(t) > 0$, $t \geq t_2$ or $z'(t) < 0$, $t \geq t_2$.

Case 1. Assume that $z'(t) > 0$, for $t \geq t_2$. Proceeding as in the proof of Theorem 3.3 and setting $\rho(t) = t$, we can obtain a contradiction with (3.31).

Case 2. Assume that $z'(t) < 0$, for $t \geq t_2$. Then there exists a finite limit

$$\lim_{t\to\infty} z(t) = l, \tag{3.61}$$

where $l \geq 0$. Next, we claim that $l = 0$. If not, then for any $\epsilon > 0$, we have $l < z(t) < l + \epsilon$, eventually. Take $0 < \epsilon < l(1-p_1)/p_1$. We calculate

$$x(t) = z(t) - p(t)x(\tau(t)) > l - p_1 z(\tau(t)) > l - p_1(l+\epsilon) = m(l+\epsilon) > mz(t), \tag{3.62}$$

where

$$m = \frac{l}{l+\epsilon} - p_1 = \frac{l(1-p_1) - \epsilon p_1}{l+\epsilon} > 0. \tag{3.63}$$

From (3.2) and (3.62), we have

$$\left(r(t)\left(-z'(t)\right)^{\alpha}\right)' \geq q(t)x^{\alpha}(\sigma(t)) \geq (ml)^{\alpha} q(t). \tag{3.64}$$

Integrating the above inequality from t_2 to t, we get

$$r(t)\left(-z'(t)\right)^{\alpha} - r(t_2)\left(-z'(t_2)\right)^{\alpha} \geq (ml)^{\alpha} \int_{t_2}^{t} q(s)\mathrm{d}s, \tag{3.65}$$

which implies

$$z'(t) \leq -ml\left[\frac{1}{r(t)}\int_{t_2}^{t} q(s)\mathrm{d}s\right]^{1/\alpha}. \tag{3.66}$$

Integrating the above inequality from t_2 to t, we have

$$z(t) \leq z(t_2) - ml\int_{t_2}^{t}\left[\frac{1}{r(s)}\int_{t_2}^{s} q(u)\mathrm{d}u\right]^{1/\alpha}\mathrm{d}s, \tag{3.67}$$

which yields $z(t) \to -\infty$; this is a contradiction. Hence, $\lim_{t\to\infty} z(t) = 0$. Note that $0 < x(t) \leq z(t)$. Then, $\lim_{t\to\infty} x(t) = 0$. The proof is complete. □

to Contents, page 1

4. Examples

In this section, we will give two examples to illustrate the main results.

Example 4.1. Consider the following linear neutral equation:

$$(x(t) + 2x(t + (2n-1)\pi))'' + x(t + (2m-1)\pi) = 0, \quad \text{for } t \geq t_0, \tag{4.1}$$

where n and m are positive integers.

Let

$$r(t) = 1, \quad p(t) = 2, \quad \tau(t) = t + (2n-1)\pi, \quad q(t) = 1, \quad \sigma(t) = t + (2m-1)\pi. \tag{4.2}$$

Hence, $Q(t) = 1$. Obviously, all the conditions of Corollary 3.5 hold. Thus by Corollary 3.5, every solution of (4.1) is oscillatory. It is easy to verify that $x(t) = \sin t$ is a solution of (4.1).

Example 4.2. Consider the following linear neutral equation:

$$\left(e^{2t}\left(x(t) + \frac{1}{2}x(t+3)\right)'\right)' + \left(e^{2t+1} + \frac{1}{2}e^{2t-2}\right)x(t+1) = 0, \quad \text{for } t \geq t_0, \tag{4.3}$$

where n and m are positive integers.

Let

$$r(t) = e^{2t}, \qquad p(t) = \frac{1}{2}, \qquad q(t) = e^{2t+1} + e^{2t-2}/2, \quad \alpha = 1. \tag{4.4}$$

Clearly, all the conditions of Theorem 3.8 hold. Thus by Theorem 3.8, every solution of (4.3) is either oscillatory or $\lim_{t\to\infty} x(t) = 0$. It is easy to verify that $x(t) = e^{-t}$ is a solution of (4.3).

Remark 4.3. Recent results cannot be applied to (4.1) and (4.3) since $\tau(t) \geq t$ and $\sigma(t) \geq t$.

Remark 4.4. Using the method given in this paper, we can get other oscillation criteria for (1.1); the details are left to the reader.

Remark 4.5. It would be interesting to find another method to study (1.1) when $\tau \circ \sigma \not\equiv \sigma \circ \tau$.

Acknowledgments

The authors sincerely thank the referees for their valuable suggestions and useful comments that have led to the present improved version of the original manuscript. This research is supported by the Natural Science Foundation of China (11071143, 60904024, 11026112), China Postdoctoral Science Foundation funded project (200902564), by Shandong Provincial Natural Science Foundation (ZR2010AL002, ZR2009AL003, Y2008A28), and also by University of Jinan Research Funds for Doctors (XBS0843).

to Contents, page 1

References

[1] J. Hale, *Theory of Functional Differential Equations*, Springer, New York, NY, USA, 2nd edition, 1977.

[2] J. Baštinec, J. Diblík, and Z. Šmarda, "Oscillation of solutions of a linear second-order discrete-delayed equation," *Advances in Difference Equations*, vol. 2010, Article ID 693867, 12 pages, 2010.

[3] J. Baštinec, L. Berezansky, J. Diblík, and Z. Šmarda, "On the critical case in oscillation for differential equations with a single delay and with several delays," *Abstract and Applied Analysis*, vol. 2010, Article ID 417869, pp. 1–20, 2010.

[4] J. Diblík, Z. Svoboda, and Z. Šmarda, "Explicit criteria for the existence of positive solutions for a scalar differential equation with variable delay in the critical case," *Computers & Mathematics with Applications. An International Journal*, vol. 56, no. 2, pp. 556–564, 2008.

[5] R. P. Agarwal, S.-L. Shieh, and C.-C. Yeh, "Oscillation criteria for second-order retarded differential equations," *Mathematical and Computer Modelling*, vol. 26, no. 4, pp. 1–11, 1997.

[6] J.-L. Chern, W.-C. Lian, and C.-C. Yeh, "Oscillation criteria for second order half-linear differential equations with functional arguments," *Publicationes Mathematicae Debrecen*, vol. 48, no. 3-4, pp. 209–216, 1996.

[7] J. Džurina and I. P. Stavroulakis, "Oscillation criteria for second-order delay differential equations," *Applied Mathematics and Computation*, vol. 140, no. 2-3, pp. 445–453, 2003.

[8] T. Kusano and N. Yoshida, "Nonoscillation theorems for a class of quasilinear differential equations of second order," *Journal of Mathematical Analysis and Applications*, vol. 189, no. 1, pp. 115–127, 1995.

[9] T. Kusano and Y. Naito, "Oscillation and nonoscillation criteria for second order quasilinear differential equations," *Acta Mathematica Hungarica*, vol. 76, no. 1-2, pp. 81–99, 1997.

[10] D. D. Mirzov, "The oscillation of the solutions of a certain system of differential equations," *Matematicheskie Zametki*, vol. 23, no. 3, pp. 401–404, 1978.

[11] Y. G. Sun and F. W. Meng, "Note on the paper of Džurina and Stavroulakis," *Applied Mathematics and Computation*, vol. 174, no. 2, pp. 1634–1641, 2006.

[12] B. Baculíková, "Oscillation criteria for second order nonlinear differential equations," *Archivum Mathematicum*, vol. 42, no. 2, pp. 141–149, 2006.

[13] Ch. G. Philos, "Oscillation theorems for linear differential equations of second order," *Archiv der Mathematik*, vol. 53, no. 5, pp. 482–492, 1989.

[14] J. Diblík, Z. Svoboda, and Z. Šmarda, "Retract principle for neutral functional differential equations," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 71, no. 12, pp. e1393–e1400, 2009.

[15] L. Erbe, T. S. Hassan, and A. Peterson, "Oscillation of second order neutral delay differential equations," *Advances in Dynamical Systems and Applications*, vol. 3, no. 1, pp. 53–71, 2008.

[16] L. Liu and Y. Bai, "New oscillation criteria for second-order nonlinear neutral delay differential equations," *Journal of Computational and Applied Mathematics*, vol. 231, no. 2, pp. 657–663, 2009.

[17] R. Xu and F. Meng, "Some new oscillation criteria for second order quasi-linear neutral delay differential equations," *Applied Mathematics and Computation*, vol. 182, no. 1, pp. 797–803, 2006.

[18] R. Xu and F. Meng, "Oscillation criteria for second order quasi-linear neutral delay differential equations," *Applied Mathematics and Computation*, vol. 192, no. 1, pp. 216–222, 2007.

[19] J.-G. Dong, "Oscillation behavior of second order nonlinear neutral differential equations with deviating arguments," *Computers & Mathematics with Applications*, vol. 59, no. 12, pp. 3710–3717, 2010.

[20] B. Baculíková and D. Lacková, "Oscillation criteria for second order retarded differential equations," *Studies of the University of Žilina. Mathematical Series*, vol. 20, no. 1, pp. 11–18, 2006.

[21] J. Jiang and X. Li, "Oscillation of second order nonlinear neutral differential equations," *Applied Mathematics and Computation*, vol. 135, no. 2-3, pp. 531–540, 2003.

[22] Q. Wang, "Oscillation theorems for first-order nonlinear neutral functional differential equations," *Computers & Mathematics with Applications*, vol. 39, no. 5-6, pp. 19–28, 2000.

[23] M. Hasanbulli and Y. V. Rogovchenko, "Oscillation criteria for second order nonlinear neutral differential equations," *Applied Mathematics and Computation*, vol. 215, no. 12, pp. 4392–4399, 2010.

[24] S. H. Saker and D. O'Regan, "New oscillation criteria for second-order neutral functional dynamic equations via the generalized Riccati substitution," *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, no. 1, pp. 423–434, 2011.

[25] B. Baculíková and J. Džurina, "Oscillation of third-order neutral differential equations," *Mathematical and Computer Modelling*, vol. 52, no. 1-2, pp. 215–226, 2010.

[26] Z. Han, T. Li, S. Sun, and Y. Sun, "Remarks on the paper, (Applied Mathematics and Computation), 2009, vol. 207, 388-396," *Applied Mathematics and Computation*, vol. 215, pp. 3998–4007, 2010.

to Contents, page 1

[27] Q. Zhang, J. Yan, and L. Gao, "Oscillation behavior of even-order nonlinear neutral differential equations with variable coefficients," *Computers & Mathematics with Applications*, vol. 59, no. 1, pp. 426–430, 2010.

[28] Z. Xu, "Oscillation theorems related to averaging technique for second order Emden-Fowler type neutral differential equations," *The Rocky Mountain Journal of Mathematics*, vol. 38, no. 2, pp. 649–667, 2008.

[29] Z. Han, T. Li, S. Sun, and W. Chen, "On the oscillation of second-order neutral delay differential equations," *Advances in Difference Equations*, vol. 2010, Article ID 289340, pp. 1–8, 2010.

[30] T. Li, Z. Han, P. Zhao, and S. Sun, "Oscillation of even-order neutral delay differential equations," *Advances in Difference Equations*, vol. 2010, Article ID 184180, pp. 1–9, 2010.

[31] Z. Han, T. Li, S. Sun, C. Zhang, and B. Han, "Oscillation criteria for a class of second-order neutral delay dynamic equations of Emden-Fowler type," *Abstract and Applied Analysis*, vol. 2011, Article ID 653689, pp. 1–26, 2011.

[32] Z. Han, T. Li, S. Sun, and C. Zhang, "An oscillation criteria for third order neutral delay differential equations," *Journal of Analytical and Applied*, vol. 16, pp. 295–303, 2010.

[33] S. Sun, T. Li, Z. Han, and Y. Sun, "Oscillation of second-order neutral functional differential equations with mixed nonlinearities," *Abstract and Applied Analysis*, vol. 2011, pp. 1–15, 2011.

[34] Z. Han, T. Li, S. Sun, and W. Chen, "Oscillation criteria for second-order nonlinear neutral delay differential equations," *Advances in Difference Equations*, vol. 2010, Article ID 763278, pp. 1–23, 2010.

[35] Z. Han, T. Li, C. Zhang, and Y. Sun, "Oscillation criteria for a certain second-order nonlinear neutral differential equations of mixed type," *Abstract and Applied Analysis*, vol. 2011, pp. 1–8, 2011.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 927690, 15 pages
doi:10.1155/2011/927690

Research Article

Oscillation of Second-Order Neutral Functional Differential Equations with Mixed Nonlinearities

Shurong Sun,[1,2] Tongxing Li,[1,3] Zhenlai Han,[1,3] and Yibing Sun[1]

[1] *School of Science, University of Jinan, Jinan, Shandong 250022, China*
[2] *Department of Mathematics and Statistics, Missouri University of Science and Technology, Rolla, MO 65409-0020, USA*
[3] *School of Control Science and Engineering, Shandong University, Jinan, Shandong 250061, China*

Correspondence should be addressed to Shurong Sun, sshrong@163.com

Received 2 September 2010; Revised 26 November 2010; Accepted 23 December 2010

Academic Editor: Miroslava Růžičková

We study the following second-order neutral functional differential equation with mixed nonlinearities $(r(t)|(u(t)+p(t)u(t-\sigma))'|^{\alpha-1}(u(t)+p(t)u(t-\sigma))')' + q_0(t)|u(\tau_0(t))|^{\alpha-1}u(\tau_0(t)) + q_1(t)|u(\tau_1(t))|^{\beta-1}u(\tau_1(t)) + q_2(t)|u(\tau_2(t))|^{\gamma-1}u(\tau_2(t)) = 0$, where $\gamma > \alpha > \beta > 0$, $\int_{t_0}^{\infty}(1/r^{1/\alpha}(t))dt < \infty$. Oscillation results for the equation are established which improve the results obtained by Sun and Meng (2006), Xu and Meng (2006), Sun and Meng (2009), and Han et al. (2010).

1. Introduction

This paper is concerned with the oscillatory behavior of the second-order neutral functional differential equation with mixed nonlinearities

$$\begin{aligned}\left(r(t)\left|(u(t)+p(t)u(t-\sigma))'\right|^{\alpha-1}(u(t)+p(t)u(t-\sigma))'\right)' + q_0(t)|u(\tau_0(t))|^{\alpha-1}u(\tau_0(t))\\ +q_1(t)|u(\tau_1(t))|^{\beta-1}u(\tau_1(t)) + q_2(t)|u(\tau_2(t))|^{\gamma-1}u(\tau_2(t)) = 0, \quad t \geq t_0,\end{aligned} \tag{1.1}$$

where $\gamma > \alpha > \beta > 0$ are constants, $r \in C^1([t_0,\infty),(0,\infty))$, $p \in C([t_0,\infty),[0,1))$, $q_i \in C([t_0,\infty),\mathbb{R})$, $i = 0,1,2$, are nonnegative, $\sigma \geq 0$ is a constant. Here, we assume that there exists $\tau \in C^1([t_0,\infty),\mathbb{R})$ such that $\tau(t) \leq \tau_i(t)$, $\tau(t) \leq t$, $\lim_{t\to\infty}\tau(t) = \infty$, and $\tau'(t) > 0$ for $t \geq t_0$.

to Contents, page 1

One of our motivations for studying (1.1) is the application of this type of equations in real word life problems. For instance, neutral delay equations appear in modeling of networks containing lossless transmission lines, in the study of vibrating masses attached to an elastic bar; see the Euler equation in some variational problems, in the theory of automatic control and in neuromechanical systems in which inertia plays an important role. We refer the reader to Hale [1] and Driver [2], and references cited therein.

Recently, there has been much research activity concerning the oscillation of second-order differential equations [3–8] and neutral delay differential equations [9–20]. For the particular case when $p(t) = 0$, (1.1) reduces to the following equation:

$$\begin{aligned}&\left(r(t)|u(t)|^{\alpha-1}u(t)\right)' + q_0(t)|u(\tau_0(t))|^{\alpha-1}u(\tau_0(t))\\&\quad + q_1(t)|u(\tau_1(t))|^{\beta-1}u(\tau_1(t)) + q_2(t)|u(\tau_2(t))|^{\gamma-1}u(\tau_2(t)) = 0, \quad t \geq t_0.\end{aligned} \tag{1.2}$$

Sun and Meng [6] established some oscillation criteria for (1.2), under the condition

$$\int_{t_0}^{\infty} \frac{1}{r^{1/\alpha}(t)} \mathrm{d}t < \infty, \tag{1.3}$$

they only obtained the sufficient condition [6, Theorem 5], which guarantees that every solution u of (1.2) oscillates or tends to zero.

Sun and Meng [7] considered the oscillation of second-order nonlinear delay differential equation

$$\left(r(t)\left|u'(t)\right|^{\alpha-1}u'(t)\right)' + q_0(t)|u(\tau_0(t))|^{\alpha-1}u(\tau_0(t)) = 0, \quad t \geq t_0 \tag{1.4}$$

and obtained some results for oscillation of (1.4), for example, under the case (1.3), they obtained some results which guarantee that every solution u of (1.4) oscillates or tends to zero, see [7, Theorem 2.2].

Xu and Meng [10] discussed the oscillation of the second-order neutral delay differential equation

$$\left(r(t)\left|(u(t) + p(t)u(t-\tau))'\right|^{\alpha-1}(u(t) + p(t)u(t-\tau))'\right)' + q(t)f(u(\sigma(t))) = 0, \quad t \geq t_0 \tag{1.5}$$

and established the sufficient condition [10, Theorem 2.3], which guarantees that every solution u of (1.5) oscillates or tends to zero.

Han et al. [11] examined the oscillation of second-order neutral delay differential equation

$$\left(r(t)\varphi(u(t))\left|(u(t) + p(t)u(t-\tau))'\right|^{\alpha-1}(u(t) + p(t)u(t-\tau))'\right)' + q(t)f(u(\sigma(t))) = 0, \quad t \geq t_0 \tag{1.6}$$

to Contents, page 1

and established some sufficient conditions for oscillation of (1.6) under the conditions (1.3) and

$$\sigma(t) \leq t - \tau. \tag{1.7}$$

The condition (1.7) can be restrictive condition, since the results cannot be applied on the equation

$$\left(e^{2t}\left(u(t) + \frac{1}{2}u(t-2)\right)'\right)' + \lambda\left(e^{2t} + \frac{1}{2}e^{2t+2}\right)u(t-1) = 0, \quad t \geq t_0. \tag{1.8}$$

The aim of this paper is to derive some sufficient conditions for the oscillation of solutions of (1.1). The paper is organized as follows. In Section 2, we establish some oscillation criteria for (1.1) under the assumption (1.3). In Section 3, we will give three examples to illustrate the main results. In Section 4, we give some conclusions for this paper.

2. Main Results

In this section, we give some new oscillation criteria for (1.1).

Below, for the sake of convenience, we denote

$$z(t) := u(t) + p(t)u(t-\sigma), \qquad R(t) := \int_{t_0}^{t} \frac{1}{r^{1/\alpha}(s)} \mathrm{d}s,$$

$$\xi(t) := r^{1/\alpha}(\tau(t)) \int_{t_1}^{t} \left(\frac{1}{r(\tau(s))}\right)^{1/\alpha} \tau'(s) \mathrm{d}s,$$

$$Q_0(t) := (1 - p(\tau_0(t)))^{\alpha} q_0(t), \qquad Q_1(t) := (1 - p(\tau_1(t)))^{\beta} q_1(t),$$

$$Q_2(t) := (1 - p(\tau_2(t)))^{\gamma} q_2(t),$$

$$\zeta_0(t) := q_0(t)\left(\frac{1}{1 + p(\rho(t))}\right)^{\alpha}, \qquad \zeta_1(t) := q_1(t)\left(\frac{1}{1 + p(\rho(t))}\right)^{\beta},$$

$$\zeta_2(t) := q_2(t)\left(\frac{1}{1 + p(\rho(t))}\right)^{\gamma},$$

$$h_0(t) := q_0(t)\left(\frac{1}{1 + p(t)}\right)^{\alpha}, \qquad h_1(t) := q_1(t)\left(\frac{1}{1 + p(t)}\right)^{\beta},$$

$$h_2(t) := q_2(t)\left(\frac{1}{1 + p(t)}\right)^{\gamma},$$

to Contents, page 1

$$\delta(t) := \int_{\rho(t)}^{\infty} \frac{1}{r^{1/\alpha}(s)} \mathrm{d}s, \qquad \pi(t) := \int_{t}^{\infty} \frac{1}{r^{1/\alpha}(s)} \mathrm{d}s, \qquad k_1 := \frac{\gamma - \beta}{\gamma - \alpha}, \qquad k_2 := \frac{\gamma - \beta}{\alpha - \beta},$$
$$\varphi(t) := q_0(t)\left(\frac{\delta(t)}{1 + p(\rho(t))}\right)^{\alpha} + q_1(t)\left(\frac{\delta(t)}{1 + p(\rho(t))}\right)^{\beta} + q_2(t)\left(\frac{\delta(t)}{1 + p(\rho(t))}\right)^{\gamma}. \tag{2.1}$$

Theorem 2.1. *Assume that* (1.3) *holds,* $p'(t) \geq 0$*, and there exists* $\rho \in C^1([t_0, \infty), \mathbb{R})$*, such that* $\rho(t) \geq t$*,* $\rho'(t) > 0$*,* $\tau_i(t) \leq \rho(t) - \sigma, i = 0, 1, 2$*. If for all sufficiently large* t_1*,*

$$\int^{\infty} \left\{ R^{\alpha}(\tau(t)) \left[Q_0(t) + [k_1 Q_1(t)]^{1/k_1} [k_2 Q_2(t)]^{1/k_2} \right] - \frac{\alpha \tau'(t) R^{\alpha-1}(\tau(t)) r^{1-1/\alpha}(\tau(t))}{\xi^{\alpha}(t)} \right\} \mathrm{d}t = \infty, \tag{2.2}$$

$$\int^{\infty} \left\{ \left[\zeta_0(t) + [k_1 \zeta_1(t)]^{1/k_1} [k_2 \zeta_2(t)]^{1/k_2} \right] \delta^{\alpha}(t) - \left(\frac{\alpha}{\alpha + 1} \right)^{\alpha+1} \frac{\rho'(t)}{\delta(t) r^{1/\alpha}(\rho(t))} \right\} \mathrm{d}t = \infty, \tag{2.3}$$

then (1.1) *is oscillatory.*

Proof. Suppose to the contrary that u is a nonoscillatory solution of (1.1). Without loss of generality, we may assume that $u(t) > 0$ for all large t. The case of $u(t) < 0$ can be considered by the same method. From (1.1) and (1.3), we can easily obtain that there exists a $t_1 \geq t_0$ such that

$$z(t) > 0, \quad z'(t) > 0, \quad \left[r(t) \left| z'(t) \right|^{\alpha-1} z'(t) \right]' \leq 0, \tag{2.4}$$

or

$$z(t) > 0, \quad z'(t) < 0, \quad \left[r(t) \left| z'(t) \right|^{\alpha-1} z'(t) \right]' \leq 0. \tag{2.5}$$

If (2.4) holds, we have

$$r(t) \left(z'(t) \right)^{\alpha} \leq r(\tau(t)) \left(z'(\tau(t)) \right)^{\alpha}, \quad t \geq t_1. \tag{2.6}$$

From the definition of z, we obtain

$$u(t) = z(t) - p(t) u(t - \sigma) \geq z(t) - p(t) z(t - \sigma) \geq \left(1 - p(t) \right) z(t). \tag{2.7}$$

Define

$$\omega(t) = R^{\alpha}(\tau(t)) \frac{r(t)(z'(t))^{\alpha}}{(z(\tau(t)))^{\alpha}}, \quad t \geq t_1. \tag{2.8}$$

to Contents, page 1

Then, $\omega(t) > 0$ for $t \geq t_1$. Noting that $z'(t) > 0$, we get $z(\tau_i(t)) \geq z(\tau(t))$ for $i = 0, 1, 2$. Thus, from (1.1), (2.7), and (2.8), it follows that

$$\begin{aligned}\omega'(t) \leq{}& \frac{\alpha\tau'(t)R^{\alpha-1}(\tau(t))}{r^{1/\alpha}(\tau(t))}\frac{r(t)(z'(t))^{\alpha}}{(z(\tau(t)))^{\alpha}} - R^{\alpha}(\tau(t))(1-p(\tau_0(t)))^{\alpha}q_0(t)\\ &- R^{\alpha}(\tau(t))\left[(1-p(\tau_1(t)))^{\beta}q_1(t)z^{\beta-\alpha}(\tau(t)) + (1-p(\tau_2(t)))^{\gamma}q_2(t)z^{\gamma-\alpha}(\tau(t))\right]\\ &- \alpha R^{\alpha}(\tau(t))\frac{r(t)(z'(t))^{\alpha}}{(z(\tau(t)))^{\alpha+1}}z'(\tau(t))\tau'(t).\end{aligned} \tag{2.9}$$

By (2.4), (2.9), and $\tau'(t) > 0$, we get

$$\begin{aligned}\omega'(t) \leq{}& \frac{\alpha\tau'(t)R^{\alpha-1}(\tau(t))}{r^{1/\alpha}(\tau(t))}\frac{r(t)(z'(t))^{\alpha}}{(z(\tau(t)))^{\alpha}} - R^{\alpha}(\tau(t))(1-p(\tau_0(t)))^{\alpha}q_0(t)\\ &- R^{\alpha}(\tau(t))\left[(1-p(\tau_1(t)))^{\beta}q_1(t)z^{\beta-\alpha}(\tau(t)) + (1-p(\tau_2(t)))^{\gamma}q_2(t)z^{\gamma-\alpha}(\tau(t))\right].\end{aligned} \tag{2.10}$$

In view of (2.4), (2.6), and (2.10), we have

$$\begin{aligned}\omega'(t) \leq{}& \frac{\alpha\tau'(t)R^{\alpha-1}(\tau(t))}{r^{1/\alpha}(\tau(t))}\frac{r(\tau(t))(z'(\tau(t)))^{\alpha}}{(z(\tau(t)))^{\alpha}} - R^{\alpha}(\tau(t))(1-p(\tau_0(t)))^{\alpha}q_0(t)\\ &- R^{\alpha}(\tau(t))\left[(1-p(\tau_1(t)))^{\beta}q_1(t)z^{\beta-\alpha}(\tau(t)) + (1-p(\tau_2(t)))^{\gamma}q_2(t)z^{\gamma-\alpha}(\tau(t))\right].\end{aligned} \tag{2.11}$$

By (2.4), we obtain

$$\begin{aligned}z(\tau(t)) &= z(\tau(t_1)) + \int_{t_1}^{t} z'(\tau(s))\tau'(s)\mathrm{d}s\\ &= z(\tau(t_1)) + \int_{t_1}^{t}\left(\frac{1}{r(\tau(s))}\right)^{1/\alpha}\left[r(\tau(s))(z'(\tau(s)))^{\alpha}\right]^{1/\alpha}\tau'(s)\mathrm{d}s\\ &\geq r^{1/\alpha}(\tau(t))z'(\tau(t))\int_{t_1}^{t}\left(\frac{1}{r(\tau(s))}\right)^{1/\alpha}\tau'(s)\mathrm{d}s,\end{aligned} \tag{2.12}$$

that is,

$$z(\tau(t)) \geq \xi(t)z'(\tau(t)). \tag{2.13}$$

Set

$$a := \left[k_1Q_1(t)z^{\beta-\alpha}(\tau(t))\right]^{1/k_1}, \quad b := \left[k_2Q_2(t)z^{\gamma-\alpha}(\tau(t))\right]^{1/k_2}, \quad p := k_1, \quad q := k_2. \tag{2.14}$$

to Contents, page 1

Using Young's inequality

$$|ab| \le \frac{1}{p}|a|^p + \frac{1}{q}|b|^q, \quad a, b \in \mathbb{R},\ p > 1,\ q > 1,\ \frac{1}{p} + \frac{1}{q} = 1, \tag{2.15}$$

we have

$$Q_1(t) z^{\beta-\alpha}(\tau(t)) + Q_2(t) z^{\gamma-\alpha}(\tau(t)) \ge [k_1 Q_1(t)]^{1/k_1} [k_2 Q_2(t)]^{1/k_2}. \tag{2.16}$$

Hence, by (2.11), (2.13), and (2.16), we obtain

$$\omega'(t) \le \frac{\alpha \tau'(t) R^{\alpha-1}(\tau(t)) r^{1-1/\alpha}(\tau(t))}{\xi^\alpha(t)} - R^\alpha(\tau(t)) \Big[Q_0(t) + [k_1 Q_1(t)]^{1/k_1} [k_2 Q_2(t)]^{1/k_2} \Big]. \tag{2.17}$$

Integrating (2.17) from t_1 to t, we get

$$0 < \omega(t) \le \omega(t_1), \tag{2.18}$$

$$-\int_{t_1}^{t} \left\{ R^\alpha(\tau(s)) \Big[Q_0(s) + [k_1 Q_1(s)]^{1/k_1} [k_2 Q_2(s)]^{1/k_2} \Big] - \frac{\alpha \tau'(s) R^{\alpha-1}(\tau(s)) r^{1-1/\alpha}(\tau(s))}{\xi^\alpha(s)} \right\} \mathrm{d}s. \tag{2.19}$$

Letting $t \to \infty$ in (2.19), we get a contradiction to (2.2). If (2.5) holds, we define the function υ by

$$\upsilon(t) = \frac{r(t)(-z'(t))^{\alpha-1} z'(t)}{z^\alpha(\rho(t))}, \quad t \ge t_1. \tag{2.20}$$

Then, $\upsilon(t) < 0$ for $t \ge t_1$. It follows from $[r(t)|z'(t)|^{\alpha-1} z'(t)]' \le 0$ that $r(t)|z'(t)|^{\alpha-1} z'(t)$ is nonincreasing. Thus, we have

$$r^{1/\alpha}(s) z'(s) \le r^{1/\alpha}(t) z'(t), \quad s \ge t. \tag{2.21}$$

Dividing (2.21) by $r^{1/\alpha}(s)$ and integrating it from $\rho(t)$ to l, we obtain

$$z(l) \le z(\rho(t)) + r^{1/\alpha}(t) z'(t) \int_{\rho(t)}^{l} \frac{\mathrm{d}s}{r^{1/\alpha}(s)}, \quad l \ge \rho(t). \tag{2.22}$$

Letting $l \to \infty$ in the above inequality, we obtain

$$0 \le z(\rho(t)) + r^{1/\alpha}(t) z'(t) \delta(t), \quad t \ge t_1, \tag{2.23}$$

to Contents, page 1

that is,

$$r^{1/\alpha}(t)\delta(t)\frac{z'(t)}{z(\rho(t))} \geq -1, \quad t \geq t_1. \tag{2.24}$$

Hence, by (2.20), we have

$$-1 \leq v(t)\delta^{\alpha}(t) \leq 0, \qquad t \geq t_1. \tag{2.25}$$

Differentiating (2.20), we get

$$v'(t) = \frac{\left(r(t)(-z'(t))^{\alpha-1}z'(t)\right)' z^{\alpha}(\rho(t)) - \alpha r(t)(-z'(t))^{\alpha-1}z'(t)z^{\alpha-1}(\rho(t))z'(\rho(t))\rho'(t)}{z^{2\alpha}(\rho(t))}, \tag{2.26}$$

by the above equality and (1.1), we obtain

$$\begin{aligned} v'(t) = &-q_0(t)\frac{u^{\alpha}(\tau_0(t))}{z^{\alpha}(\rho(t))} - q_1(t)\frac{u^{\beta}(\tau_1(t))}{z^{\alpha}(\rho(t))} - q_2(t)\frac{u^{\gamma}(\tau_2(t))}{z^{\alpha}(\rho(t))} \\ &- \frac{\alpha r(t)(-z'(t))^{\alpha-1}z'(t)z^{\alpha-1}(\rho(t))z'(\rho(t))\rho'(t)}{z^{2\alpha}(\rho(t))}. \end{aligned} \tag{2.27}$$

Noticing that $p'(t) \geq 0$, from [10, Theorem 2.3], we see that $u'(t) \leq 0$ for $t \geq t_1$, so by $\tau_i(t) \leq \rho(t) - \sigma$, $i = 0, 1, 2$, we have

$$\begin{aligned} \frac{u^{\alpha}(\tau_0(t))}{z^{\alpha}(\rho(t))} &= \left(\frac{u(\tau_0(t))}{u(\rho(t)) + p(\rho(t))u(\rho(t) - \sigma)}\right)^{\alpha} \\ &= \left(\frac{1}{(u(\rho(t))/u(\tau_0(t))) + p(\rho(t))(u(\rho(t) - \sigma)/u(\tau_0(t)))}\right)^{\alpha} \\ &\geq \left(\frac{1}{1 + p(\rho(t))}\right)^{\alpha}, \\ \frac{u^{\beta}(\tau_1(t))}{z^{\alpha}(\rho(t))} &= \left(\frac{u(\tau_1(t))}{u(\rho(t)) + p(\rho(t))u(\rho(t) - \sigma)}\right)^{\beta} z^{\beta-\alpha}(\rho(t)) \\ &= \left(\frac{1}{(u(\rho(t))/u(\tau_1(t))) + p(\rho(t))(u(\rho(t) - \sigma)/u(\tau_1(t)))}\right)^{\beta} z^{\beta-\alpha}(\rho(t)) \\ &\geq \left(\frac{1}{1 + p(\rho(t))}\right)^{\beta} z^{\beta-\alpha}(\rho(t)), \end{aligned}$$

to Contents, page 1

$$\begin{aligned}\left(u^{\gamma}(\tau_2(t))/z^{\alpha}(\rho(t))\right) &= \left(\frac{u(\tau_2(t))}{u(\rho(t))+p(\rho(t))u(\rho(t)-\sigma)}\right)^{\gamma} z^{\gamma-\alpha}(\rho(t)) \\ &= \left(\frac{1}{(u(\rho(t))/u(\tau_2(t)))+p(\rho(t))(u(\rho(t)-\sigma)/u(\tau_2(t)))}\right)^{\gamma} z^{\gamma-\alpha}(\rho(t)) \\ &\geq \left(\frac{1}{1+p(\rho(t))}\right)^{\gamma} z^{\gamma-\alpha}(\rho(t)).\end{aligned} \tag{2.28}$$

On the other hand, from $(r(t)(-z'(t))^{\alpha-1}z'(t))' \leq 0$, $\rho(t) \geq t$, we obtain

$$z'(\rho(t)) \leq \frac{r^{1/\alpha}(t)}{r^{1/\alpha}(\rho(t))} z'(t). \tag{2.29}$$

Thus, by (2.20) and (2.27), we get

$$v'(t) \leq -\left[\zeta_0(t)+\zeta_1(t)z^{\beta-\alpha}(\rho(t))+\zeta_2(t)z^{\gamma-\alpha}(\rho(t))\right] - \frac{\alpha\rho'(t)}{r^{1/\alpha}(\rho(t))}(-v(t))^{(\alpha+1)/\alpha}. \tag{2.30}$$

Set

$$a := \left[k_1\zeta_1(t)z^{\beta-\alpha}(\rho(t))\right]^{1/k_1}, \quad b := \left[k_2\zeta_2(t)z^{\gamma-\alpha}(\rho(t))\right]^{1/k_2},\ p := k_1,\ q := k_2. \tag{2.31}$$

Using Young's inequality (2.15), we obtain

$$\zeta_1(t)z^{\beta-\alpha}(\rho(t))+\zeta_2(t)z^{\gamma-\alpha}(\rho(t)) \geq [k_1\zeta_1(t)]^{1/k_1}[k_2\zeta_2(t)]^{1/k_2}. \tag{2.32}$$

Hence, from (2.30), we have

$$v'(t) \leq -\left[\zeta_0(t)+[k_1\zeta_1(t)]^{1/k_1}[k_2\zeta_2(t)]^{1/k_2}\right] - \frac{\alpha\rho'(t)}{r^{1/\alpha}(\rho(t))}(-v(t))^{(\alpha+1)/\alpha}, \tag{2.33}$$

that is,

$$v'(t) + \left[\zeta_0(t)+[k_1\zeta_1(t)]^{1/k_1}[k_2\zeta_2(t)]^{1/k_2}\right] + \frac{\alpha\rho'(t)}{r^{1/\alpha}(\rho(t))}(-v(t))^{(\alpha+1)/\alpha} \leq 0, \quad t \geq t_1. \tag{2.34}$$

to Contents, page 1

Multiplying (2.34) by $\delta^{\alpha}(t)$ and integrating it from t_1 to t implies that

$$\delta^{\alpha}(t)v(t) - \delta^{\alpha}(t_1)v(t_1) + \alpha\int_{t_1}^{t} r^{-1/\alpha}(\rho(s))\rho'(s)\delta^{\alpha-1}(s)v(s)\mathrm{d}s$$
$$+ \int_{t_1}^{t}\left[\zeta_0(s) + [k_1\zeta_1(s)]^{1/k_1}[k_2\zeta_2(s)]^{1/k_2}\right]\delta^{\alpha}(s)\mathrm{d}s \tag{2.35}$$
$$+ \alpha\int_{t_1}^{t}\frac{\delta^{\alpha}(s)\rho'(s)}{r^{1/\alpha}(\rho(s))}(-v(s))^{(\alpha+1)/\alpha}\mathrm{d}s \le 0.$$

Set $p := (\alpha+1)/\alpha$, $q := \alpha+1$, and

$$a := (\alpha+1)^{\alpha/(\alpha+1)}\delta^{\alpha^2/(\alpha+1)}(t)v(t), \qquad b := \frac{\alpha}{(\alpha+1)^{\alpha/(\alpha+1)}}\delta^{-1/(\alpha+1)}(t). \tag{2.36}$$

Using Young's inequality (2.15), we get

$$-\alpha\delta^{\alpha-1}(t)v(t) \le \alpha\delta^{\alpha}(t)(-v(t))^{(\alpha+1)/\alpha} + \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1}\frac{1}{\delta(t)}. \tag{2.37}$$

Thus,

$$-\frac{\alpha\rho'(t)\delta^{\alpha-1}(t)v(t)}{r^{1/\alpha}(\rho(t))} \le \alpha\rho'(t)\frac{\delta^{\alpha}(t)(-v(t))^{(\alpha+1)/\alpha}}{r^{1/\alpha}(\rho(t))} + \rho'(t)\left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1}\frac{1}{\delta(t)r^{1/\alpha}(\rho(t))}. \tag{2.38}$$

Therefore, (2.35) yields

$$\delta^{\alpha}(t)v(t) \le \delta^{\alpha}(t_1)v(t_1),$$
$$-\int_{t_1}^{t}\left\{\left[\zeta_0(s) + [k_1\zeta_1(s)]^{1/k_1}[k_2\zeta_2(s)]^{1/k_2}\right]\delta^{\alpha}(s) - \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1}\frac{\rho'(s)}{\delta(s)r^{1/\alpha}(\rho(s))}\right\}\mathrm{d}s. \tag{2.39}$$

Letting $t \to \infty$ in the above inequality, by (2.3), we get a contradiction with (2.25). This completes the proof of Theorem 2.1. □

From Theorem 2.1, when $\rho(t) = t$, we have the following result.

Corollary 2.2. *Assume that* (1.3) *holds,* $p'(t) \ge 0$, *and* $\tau_i(t) \le t - \sigma$, $i = 0, 1, 2$. *If for all sufficiently large* t_1 *such that* (2.2) *holds and*

$$\int^{\infty}\left\{\left[h_0(t) + [k_1h_1(t)]^{1/k_1}[k_2h_2(t)]^{1/k_2}\right]\pi^{\alpha}(t) - \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1}\frac{1}{\pi(t)r^{1/\alpha}(t)}\right\}\mathrm{d}t = \infty, \tag{2.40}$$

then (1.1) *is oscillatory.*

to Contents, page 1

Theorem 2.3. *Assume that* (1.3) *holds,* $p'(t) \geq 0$, *and there exists* $\rho \in C^1([t_0, \infty), \mathbb{R})$, *such that* $\rho(t) \geq t$, $\rho'(t) > 0$, $\tau_i(t) \leq \rho(t) - \sigma$, $i = 0, 1, 2$. *If for all sufficiently large* t_1 *such that* (2.2) *holds and*

$$\int^{\infty} \left[\zeta_0(t) + [k_1\zeta_1(t)]^{1/k_1}[k_2\zeta_2(t)]^{1/k_2}\right]\delta^{\alpha+1}(t)\mathrm{d}t = \infty, \tag{2.41}$$

then (1.1) *is oscillatory.*

Proof. Suppose to the contrary that u is a nonoscillatory solution of (1.1). Without loss of generality, we may assume that $u(t) > 0$ for all large t. The case of $u(t) < 0$ can be considered by the same method. From (1.1) and (1.3), we can easily obtain that there exists a $t_1 \geq t_0$ such that (2.4) or (2.5) holds.

If (2.4) holds, proceeding as in the proof of Theorem 2.1, we obtain a contradiction with (2.2).

If (2.5) holds, we proceed as in the proof of Theorem 2.1, then we get (2.25) and (2.34). Multiplying (2.34) by $\delta^{\alpha+1}(t)$ and integrating it from t_1 to t implies that

$$\begin{aligned}
&\delta^{\alpha+1}(t)v(t) - \delta^{\alpha+1}(t_1)v(t_1) + (\alpha+1)\int_{t_1}^{t} r^{-1/\alpha}(\rho(s))\rho'(s)\delta^{\alpha}(s)v(s)\mathrm{d}s \\
&\quad + \int_{t_1}^{t}\left[\zeta_0(s) + [k_1\zeta_1(s)]^{1/k_1}[k_2\zeta_2(s)]^{1/k_2}\right]\delta^{\alpha+1}(s)\mathrm{d}s \\
&\quad + \alpha\int_{t_1}^{t}\frac{\delta^{\alpha+1}(s)\rho'(s)}{r^{1/\alpha}(\rho(s))}(-v(s))^{(\alpha+1)/\alpha}\mathrm{d}s \leq 0.
\end{aligned} \tag{2.42}$$

In view of (2.25), we have $-v(t)\delta^{\alpha+1}(t) \leq \delta(t) < \infty$, $t \to \infty$. From (1.3), we get

$$\begin{aligned}
&\int_{t_1}^{t} -r^{-1/\alpha}(\rho(s))\rho'(s)\delta^{\alpha}(s)v(s)\mathrm{d}s \leq \int_{t_1}^{t} r^{-1/\alpha}(\rho(s))\rho'(s)\mathrm{d}s = \int_{\rho(t_1)}^{\rho(t)} r^{-1/\alpha}(u)\mathrm{d}u < \infty, \quad t \longrightarrow \infty, \\
&\int_{t_1}^{t}\frac{\delta^{\alpha+1}(s)\rho'(s)}{r^{1/\alpha}(\rho(s))}(-v(s))^{(\alpha+1)/\alpha}\mathrm{d}s \leq \int_{\rho(t_1)}^{\rho(t)} r^{-1/\alpha}(u)\mathrm{d}u < \infty, \quad t \longrightarrow \infty.
\end{aligned} \tag{2.43}$$

Letting $t \to \infty$ in (2.42) and using the last inequalities, we obtain

$$\int^{\infty} \left[\zeta_0(t) + [k_1\zeta_1(t)]^{1/k_1}[k_2\zeta_2(t)]^{1/k_2}\right]\delta^{\alpha+1}(t)\mathrm{d}t < \infty, \tag{2.44}$$

which contradicts (2.41). This completes the proof of Theorem 2.3. □

From Theorem 2.3, when $\rho(t) = t$, we have the following result.

to Contents, page 1

Corollary 2.4. *Assume that* (1.3) *holds,* $p'(t) \geq 0$*,* $\tau_i(t) \leq t - \sigma$*,* $i = 0, 1, 2$*. If for all sufficiently large* t_1 *such that* (2.2) *holds and*

$$\int^{\infty} \left[h_0(t) + [k_1 h_1(t)]^{1/k_1} [k_2 h_2(t)]^{1/k_2} \right] \pi^{\alpha+1}(t) \mathrm{d}t = \infty, \tag{2.45}$$

then (1.1) *is oscillatory.*

Theorem 2.5. *Assume that* (1.3) *holds,* $p'(t) \geq 0$*, and there exists* $\rho \in C^1([t_0, \infty), \mathbb{R})$*, such that* $\rho(t) \geq t$*,* $\rho'(t) > 0$*,* $\tau_i(t) \leq \rho(t) - \sigma$*,* $i = 0, 1, 2$*. If for all sufficiently large* t_1 *such that* (2.2) *holds and*

$$\int_{t_1}^{\infty} r^{-1/\alpha}(v) \left[\int_{t_1}^{v} \varphi(u) \mathrm{d}u \right]^{1/\alpha} \mathrm{d}v = \infty, \tag{2.46}$$

then (1.1) *is oscillatory.*

Proof. Suppose to the contrary that u is a nonoscillatory solution of (1.1). Without loss of generality, we may assume that $u(t) > 0$ for all large t. The case of $u(t) < 0$ can be considered by the same method. From (1.1) and (1.3), we can easily obtain that there exists a $t_1 \geq t_0$ such that (2.4) or (2.5) holds.

If (2.4) holds, proceeding as in the proof of Theorem 2.1, we obtain a contradiction with (2.2).

If (2.5) holds, we proceed as in the proof of Theorem 2.1, and we get (2.21). Dividing (2.21) by $r^{1/\alpha}(s)$ and integrating it from $\rho(t)$ to l, letting $l \to \infty$, yields

$$z(\rho(t)) \geq -r^{1/\alpha}(t) z'(t) \int_{\rho(t)}^{\infty} r^{-1/\alpha}(s) \mathrm{d}s = -r^{1/\alpha}(t) z'(t) \delta(t) \geq -r^{1/\alpha}(t_1) z'(t_1) \delta(t) := a\delta(t). \tag{2.47}$$

By (1.1), we have

$$\left(r(t) \left(-z'(t) \right)^{\alpha} \right)' = q_0(t) u^{\alpha}(\tau_0(t)) + q_1(t) u^{\beta}(\tau_1(t)) + q_2(t) u^{\gamma}(\tau_2(t)). \tag{2.48}$$

Noticing that $p'(t) \geq 0$, from [10, Theorem 2.3], we see that $u'(t) \leq 0$ for $t \geq t_1$, so by $\tau_i(t) \leq \rho(t) - \sigma$, $i = 0, 1, 2$, we get

$$\begin{aligned} \frac{u(\tau_i(t))}{z(\rho(t))} &= \frac{u(\tau_i(t))}{u(\rho(t)) + p(\rho(t)) u(\rho(t) - \sigma)} \\ &= \frac{1}{(u(\rho(t))/u(\tau_i(t))) + p(\rho(t))(u(\rho(t) - \sigma)/u(\tau_i(t)))} \geq \frac{1}{1 + p(\rho(t))}. \end{aligned} \tag{2.49}$$

Hence, we obtain

$$\left(r(t) \left(-z'(t) \right)^{\alpha} \right)' \geq b\varphi(t), \tag{2.50}$$

to Contents, page 1

where $b = \min\{a^{\alpha}, a^{\beta}, a^{\gamma}\}$. Integrating the above inequality from t_1 to t, we have

$$r(t)\left(-z'(t)\right)^{\alpha} \geq r(t_1)\left(-z'(t_1)\right)^{\alpha} + b\int_{t_1}^{t} \varphi(u)\mathrm{d}u \geq b\int_{t_1}^{t} \varphi(u)\mathrm{d}u. \tag{2.51}$$

Integrating the above inequality from t_1 to t, we obtain

$$z(t_1) - z(t) \geq b^{1/\alpha}\int_{t_1}^{t} r^{-1/\alpha}(v)\left[\int_{t_1}^{v} \varphi(u)\mathrm{d}u\right]^{1/\alpha}\mathrm{d}v, \tag{2.52}$$

which contradicts (2.46). This completes the proof of Theorem 2.5. □

3. Examples

In this section, three examples are worked out to illustrate the main results.

Example 3.1. Consider the second-order neutral delay differential equation (1.8), where $\lambda > 0$ is a constant.

Let $r(t) = \mathrm{e}^{2t}$, $p(t) = 1/2$, $\sigma = 2$, $q_0(t) = \lambda(2\mathrm{e}^{2t} + \mathrm{e}^{2t+2})/2$, $\alpha = 1$, $\tau_0(t) = t - 1$, $q_1(t) = q_2(t) = 0$, and $\tau(t) = \tau_0(t)$, then

$$R(t) = \int_{t_0}^{t} \frac{1}{r^{1/\alpha}(s)}\mathrm{d}s = \frac{(\mathrm{e}^{-2t_0} - \mathrm{e}^{-2t})}{2},$$

$$\xi(t) = r^{1/\alpha}(\tau(t))\int_{t_1}^{t}\left(\frac{1}{r(\tau(s))}\right)^{1/\alpha}\tau'(s)\mathrm{d}s = \frac{(\mathrm{e}^{2(t-t_1)} - 1)}{2}, \tag{3.1}$$

$$Q_0(t) = \frac{q_0(t)}{2} = \frac{\lambda(2\mathrm{e}^{2t} + \mathrm{e}^{2t+2})}{4}, \qquad \zeta_0(t) = \frac{2q_0(t)}{3} = \frac{\lambda(2\mathrm{e}^{2t} + \mathrm{e}^{2t+2})}{3}.$$

Setting $\rho(t) = t + 1$, we have $\tau_0(t) = t - 1 \leq \rho(t) - \sigma$, $\delta(t) = \mathrm{e}^{-2t-2}/2$. Therefore, for all sufficiently large t_1,

$$\begin{gathered}\int^{\infty}\left\{R^{\alpha}(\tau(t))\left[Q_0(t) + [k_1Q_1(t)]^{1/k_1}[k_2Q_2(t)]^{1/k_2}\right] - \frac{\alpha\tau'(t)R^{\alpha-1}(\tau(t))r^{1-1/\alpha}(\tau(t))}{\xi^{\alpha}(t)}\right\}\mathrm{d}t = \infty,\\
\int^{\infty}\left\{\left[\zeta_0(t) + [k_1\zeta_1(t)]^{1/k_1}[k_2\zeta_2(t)]^{1/k_2}\right]\delta^{\alpha}(t) - \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1}\frac{\rho'(t)}{\delta(t)r^{1/\alpha}(\rho(t))}\right\}\mathrm{d}t\\
= \int^{\infty}\frac{\lambda(2\mathrm{e}^{-2}+1) - 3}{6}\mathrm{d}t = \infty\end{gathered} \tag{3.2}$$

if $\lambda > 3/(2\mathrm{e}^{-2} + 1)$. Hence, by Theorem 2.1, (1.8) is oscillatory when $\lambda > 3/(2\mathrm{e}^{-2} + 1)$.

to Contents, page 1

Note that [11, Theorem 2.1] and [11, Theorem 2.2] cannot be applied in (1.8), since $\tau_0(t) > t-2$. On the other hand, applying [11, Theorem 3.2] to that (1.8), we obtain that (1.8) is oscillatory if $\lambda > 3/(e^{-2}+2e^{-4})$. So our results improve the results in [11].

Example 3.2. Consider the second-order neutral delay differential equation

$$\left(e^t\left(u(t)+\frac{1}{2}u\left(t-\frac{\pi}{4}\right)\right)'\right)' + 12\sqrt{65}e^t u\left(t-\frac{1}{8}\arcsin\frac{\sqrt{65}}{65}\right) = 0, \quad t \ge t_0. \tag{3.3}$$

Let $r(t) = e^t$, $p(t) = 1/2$, $\sigma = \pi/4$, $q_0(t) = 12\sqrt{65}e^t$, $q_1(t) = q_2(t) = 0$, $\alpha = 1$, $\tau_0(t) = t - (\arcsin\sqrt{65}/65)/8$, $\rho(t) = t+\pi/4$, and $\tau(t) = t-\pi/4$, then

$$\begin{gathered} R(t) = \int_{t_0}^{t}\frac{1}{r^{1/\alpha}(s)}ds = e^{-t_0}-e^{-t}, \qquad \xi(t) = r^{1/\alpha}(\tau(t))\int_{t_1}^{t}\left(\frac{1}{r(\tau(s))}\right)^{1/\alpha}\tau'(s)ds = e^{t-t_1}-1, \\ Q_0(t) = \frac{q_0(t)}{2} = 6\sqrt{65}e^t, \qquad \zeta_0(t) = \frac{2q_0(t)}{3} = 8\sqrt{65}e^t, \qquad \delta(t) = e^{-t-\pi/4}. \end{gathered} \tag{3.4}$$

Therefore, for all sufficiently large t_1,

$$\begin{aligned} &\int^{\infty}\left\{R^{\alpha}(\tau(t))\left[Q_0(t)+[k_1Q_1(t)]^{1/k_1}[k_2Q_2(t)]^{1/k_2}\right] - \frac{\alpha\tau'(t)R^{\alpha-1}(\tau(t))r^{1-1/\alpha}(\tau(t))}{\xi^{\alpha}(t)}\right\}dt = \infty, \\ &\int^{\infty}\left\{\left[\zeta_0(t)+[k_1\zeta_1(t)]^{1/k_1}[k_2\zeta_2(t)]^{1/k_2}\right]\delta^{\alpha}(t) - \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1}\frac{\rho'(t)}{\delta(t)r^{1/\alpha}(\rho(t))}\right\}dt \\ &\quad = \int^{\infty}\left(8\sqrt{65}e^{-\pi/4}-\frac{1}{4}\right)dt = \infty. \end{aligned} \tag{3.5}$$

Hence, by Theorem 2.1, (3.3) oscillates. For example, $u(t) = \sin 8t$ is a solution of (3.3).

Example 3.3. Consider the second-order neutral differential equation

$$\left(e^t z'(t)\right)' + e^{2\lambda_* t}u(\lambda_0 t) + q_1(t)u^{1/3}(\lambda_1 t) + q_2(t)u^{5/3}(\lambda_2 t) = 0, \quad t \ge t_0, \tag{3.6}$$

where $z(t) = u(t) + u(t-1)/2$, $\lambda_i > 0$ for $i = 0,1,2$, are constants, $q_1(t) > 0$, $q_2(t) > 0$ for $t \ge t_0$.

Let $r(t) = e^t$, $\sigma = 1$, $q_0(t) = e^{2\lambda_* t}$, $\lambda_* = \max\{\lambda_0, \lambda_1, \lambda_2\}$, $\tau_i(t) = \lambda_i t$, $\tau(t) = \lambda t$, $0 < \lambda < \min\{\lambda_0, \lambda_1, \lambda_2, 1\}$, $\rho(t) = \lambda_* t + 1$, $\alpha = 1$, $\beta = 1/3$, and $\gamma = 5/3$, then $k_1 = k_2 = 2$,

$$\begin{gathered} R(t) = \int_{t_0}^{t}\frac{1}{r^{1/\alpha}(s)}ds = e^{-t_0}-e^{-t}, \\ \xi(t) = r^{1/\alpha}(\tau(t))\int_{t_1}^{t}\left(\frac{1}{r(\tau(s))}\right)^{1/\alpha}\tau'(s)ds = e^{\lambda(t-t_1)}-1, \qquad \delta(t) = e^{-\lambda_* t-1}. \end{gathered} \tag{3.7}$$

to Contents, page 1

It is easy to see that (2.2) and (2.41) hold for all sufficiently large t_1. Hence, by Theorem 2.3, (3.6) is oscillatory.

4. Conclusions

In this paper, we consider the oscillatory behavior of second-order neutral functional differential equation (1.1). Our results can be applied to the case when $\tau_i(t) > t$, $i = 0, 1, 2$; these results improve the results given in [6, 7, 10, 11].

Acknowledgments

The authors sincerely thank the reviewers for their valuable suggestions and useful comments that have led to the present improved version of the original paper. This research is supported by the Natural Science Foundation of China (nos. 11071143, 60904024, 11026112), China Postdoctoral Science Foundation funded Project (no. 200902564), the Natural Science Foundation of Shandong (nos. ZR2010AL002, ZR2009AL003, Y2008A28), and also the University of Jinan Research Funds for Doctors (no. XBS0843).

References

[1] J. K. Hale, *Theory of Functional Differential Equations*, Springer, New York, NY, USA, 2nd edition, 1977.

[2] R. D. Driver, "A mixed neutral system," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 8, no. 2, pp. 155–158, 1984.

[3] R. P. Agarwal, S.-L. Shieh, and C.-C. Yeh, "Oscillation criteria for second-order retarded differential equations," *Mathematical and Computer Modelling*, vol. 26, no. 4, pp. 1–11, 1997.

[4] B. Baculíková, "Oscillation criteria for second order nonlinear differential equations," *Archivum Mathematicum*, vol. 42, no. 2, pp. 141–149, 2006.

[5] J. Džurina and I. P. Stavroulakis, "Oscillation criteria for second-order delay differential equations," *Applied Mathematics and Computation*, vol. 140, no. 2-3, pp. 445–453, 2003.

[6] Y. G. Sun and F. W. Meng, "Oscillation of second-order delay differential equations with mixed nonlinearities," *Applied Mathematics and Computation*, vol. 207, no. 1, pp. 135–139, 2009.

[7] Y. G. Sun and F. W. Meng, "Note on the paper of Džurina and Stavroulakis," *Applied Mathematics and Computation*, vol. 174, no. 2, pp. 1634–1641, 2006.

[8] S. Sun, Z. Han, P. Zhao, and C. Zhang, "Oscillation for a class of second-order Emden-Fowler delay dynamic equations on time scales," *Advances in Difference Equations*, vol. 2010, Article ID 642356, 15 pages, 2010.

[9] B. Baculíková and D. Lacková, "Oscillation criteria for second order retarded differential equations," *Studies of the University of Žilina. Mathematical Series*, vol. 20, no. 1, pp. 11–18, 2006.

[10] R. Xu and F. Meng, "Some new oscillation criteria for second order quasi-linear neutral delay differential equations," *Applied Mathematics and Computation*, vol. 182, no. 1, pp. 797–803, 2006.

[11] Z. Han, T. Li, S. Sun, and Y. Sun, "Remarks on the paper [Appl. Math. Comput. 207 (2009) 388–396]," *Applied Mathematics and Computation*, vol. 215, no. 11, pp. 3998–4007, 2010.

[12] Z. Han, T. Li, S. Sun, and W. Chen, "On the oscillation of second-order neutral delay differential equations," *Advances in Difference Equations*, vol. 2010, Article ID 289340, 8 pages, 2010.

[13] R. Xu and F. Meng, "Oscillation criteria for second order quasi-linear neutral delay differential equations," *Applied Mathematics and Computation*, vol. 192, no. 1, pp. 216–222, 2007.

[14] L. Liu and Y. Bai, "New oscillation criteria for second-order nonlinear neutral delay differential equations," *Journal of Computational and Applied Mathematics*, vol. 231, no. 2, pp. 657–663, 2009.

[15] J.-G. Dong, "Oscillation behavior of second order nonlinear neutral differential equations with deviating arguments," *Computers & Mathematics with Applications*, vol. 59, no. 12, pp. 3710–3717, 2010.

[16] Z. Han, T. Li, S. Sun, C. Zhang, and B. Han, " Oscillation criteria for a class of second order neutral delay dynamic equations of Emden-Fowler type,," *Abstract and Applied Analysis*, vol. 2011, pp. 1–26, 2011.

to Contents, page 1

[17] T. Li, Z. Han, P. Zhao, and S. Sun, "Oscillation of even-order neutral delay differential equations," *Advances in Difference Equations*, vol. 2010, Article ID 184180, 9 pages, 2010.
[18] Z. Han, T. Li, S. Sun, and C. Zhang, "An oscillation criteria for third order neutral delay differential equations," *Journal of Applied Analysis*, vol. 16, pp. 295–303, 2010.
[19] Y. Sun, Z. Han, T. Li, and G. Zhang, "Oscillation criteria for second-order quasilinear neutral delay dynamic equations on time scales," *Advances in Difference Equations*, vol. 2010, Article ID 512437, 14 pages, 2010.
[20] Z. Han, T. Li, S. Sun, and W. Chen, "Oscillation criteria for second-order nonlinear neutral delay differential equations," *Advances in Difference Equations*, vol. 2010, Article ID 763278, 23 pages, 2010.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 458275, 11 pages
doi:10.1155/2011/458275

Research Article

Oscillation of Second-Order Sublinear Impulsive Differential Equations

A. Zafer

Department of Mathematics, Middle East Technical University, 06800 Ankara, Turkey

Correspondence should be addressed to A. Zafer, zafer@metu.edu.tr

Received 25 January 2011; Accepted 27 February 2011

Academic Editor: Josef Diblík

Oscillation criteria obtained by Kusano and Onose (1973) and by Belohorec (1969) are extended to second-order sublinear impulsive differential equations of Emden-Fowler type: $x''(t) + p(t)|x(\tau(t))|^{\alpha-1}x(\tau(t)) = 0$, $t \neq \theta_k$; $\Delta x'(t)|_{t=\theta_k} + q_k|x(\tau(\theta_k))|^{\alpha-1}x(\tau(\theta_k)) = 0$; $\Delta x(t)|_{t=\theta_k} = 0$, $(0 < \alpha < 1)$ by considering the cases $\tau(t) \leq t$ and $\tau(t) = t$, respectively. Examples are inserted to show how impulsive perturbations greatly affect the oscillation behavior of the solutions.

1. Introduction

We deal with second-order sublinear impulsive differential equations of the form

$$\begin{gathered} x''(t) + p(t)|x(\tau(t))|^{\alpha-1}x(\tau(t)) = 0, \quad t \neq \theta_k, \\ \Delta x'(t)\big|_{t=\theta_k} + q_k|x(\tau(\theta_k))|^{\alpha-1}x(\tau(\theta_k)) = 0, \\ \Delta x(t)|_{t=\theta_k} = 0, \end{gathered} \tag{1.1}$$

where $0 < \alpha < 1$, $t \geq t_0$, and $k \geq k_0$ for some $t_0 \in \mathbb{R}_+$ and $k_0 \in \mathbb{N}$, $\{\theta_k\}$ is a strictly increasing unbounded sequence of positive real numbers,

$$\Delta z(t)|_{t=\theta} := z(\theta^+) - z(\theta^-), \quad z(\theta^\mp) := \lim_{t\to\theta^\mp} z(t). \tag{1.2}$$

Let $\mathrm{PLC}(J, R)$ denote the set of all real-valued functions u defined on J such that u is continuous for all $t \in J$ except possibly at $t = \theta_k$ where $u(\theta_k^\pm)$ exists and $u(\theta_k) := u(\theta_k^-)$.

to Contents, page 1

We assume in the sequel that

(a) $p \in \mathrm{PLC}([t_0, \infty), \mathbb{R})$,

(b) $\{q_k\}$ is a sequence of real numbers,

(c) $\tau \in C([t_0, \infty), \mathbb{R}_+)$, $\tau(t) \le t$, $\lim_{t\to\infty}\tau(t) = \infty$.

By a solution of (1.1) on an interval $J \subset [t_0, \infty)$, we mean a function $x(t)$ which is defined on J such that $x, x', x'' \in \mathrm{PLC}(J)$ and which satisfies (1.1). Because of the requirement $\Delta x(t)|_{t=\theta_k} = 0$ every solution of (1.1) is necessarily continuous.

As usual we assume that (1.1) has solutions which are nontrivial for all large t. Such a solution of (1.1) is called oscillatory if it has no last zero and nonoscillatory otherwise.

In case there is no impulse, (1.1) reduces to Emden-Fowler equation with delay

$$x''(t) + p(t)|x(\tau(t))|^{\alpha-1}x(\tau(t)) = 0, \quad 0 < \alpha < 1, \tag{1.3}$$

and without delay

$$x'' + p(t)|x|^{\alpha-1}x = 0, \quad 0 < \alpha < 1. \tag{1.4}$$

The problem of oscillation of solutions of (1.3) and (1.4) has been considered by many authors. Kusano and Onose [1] see also [2, 3] proved the following necessary and sufficient condition for oscillation of (1.3).

Theorem 1.1. *If $p(t) \ge 0$, then a necessary and sufficient condition for every solution of (1.3) to be oscillatory is that*

$$\int^{\infty} [\tau(t)]^{\alpha} p(t)\,dt = \infty. \tag{1.5}$$

The condition $p(t) \ge 0$ is required only for the sufficiency part, and no similar criteria is available for $p(t)$ changing sign, except in the case $\tau(t) = t$. Without imposing a sign condition on $p(t)$, Belohorec [4] obtained the following sufficient condition for oscillation of (1.4).

Theorem 1.2. *If*

$$\int^{\infty} t^{\beta} p(t)\,dt = \infty \tag{1.6}$$

for some $\beta \in [0, \alpha]$, then every solution of (1.4) is oscillatory.

Compared to the large body of papers on oscillation of differential equations, there is only little known about the oscillation of impulsive differential equations; see [5–7] for equations with delay and [8–13] for equations without delay. For some applications of such equations, we may refer to [14–18]. The books [19, 20] are good sources for a general theory of impulsive differential equations.

The object of this paper is to extend Theorems 1.1 and 1.2 to impulsive differential equations of the form (1.1). The results show that the impulsive perturbations may greatly

to Contents, page 1

change the oscillatory behavior of the solutions. A nonoscillatory solution of (1.3) or (1.4) may become oscillatory under impulsive perturbations.

The following two lemmas are crucial in the proof of our main theorems. The first lemma is contained in [21] and the second one is extracted from [22].

Lemma 1.3. *If each A_i is continuous on $[a,b]$, then*

$$\int_a^b \sum_{s\le\theta_i<b} A_i(s)ds = \sum_{a\le\theta_i<b} \int_a^{\theta_i} A_i(s)ds. \tag{1.7}$$

Lemma 1.4. Fix $J=[a,b]$, *let* $u,\lambda\in C(J,\mathbb{R}_+)$, $h\in C(\mathbb{R}_+,\mathbb{R}_+)$, *and* $c\in\mathbb{R}_+$, *and let* $\{\lambda_k\}$ *a sequence of positive real numbers. If* $u(J)\subset I\subset\mathbb{R}_+$ *and*

$$u(t)\le c+\int_a^t \lambda(s)h(u(s))ds+\sum_{a<\theta_k<t}\lambda_k h(u(\theta_k)), \quad t\in J, \tag{1.8}$$

then

$$u(t)\le G^{-1}\left\{G(c)+\int_a^t \lambda(s)ds+\sum_{a<\theta_k<t}\lambda_k\right\}, \quad t\in[a,\beta), \tag{1.9}$$

where

$$\begin{gathered} G(u)=\int_{u_0}^{u}\frac{dx}{h(x)}, \quad u,u_0\in I, \\ \beta=\sup\left\{\nu\in J: G(c)+\int_a^t \lambda(s)ds+\sum_{a<\theta_k<t}\lambda_k\in G(I),\ a\le t\le\nu\right\}. \end{gathered} \tag{1.10}$$

2. The Main Results

We first establish a necessary and sufficient condition for oscillation of solutions of (1.1) when $\tau(t)\le t$.

Theorem 2.1. *If*

$$\int^{\infty}[\tau(t)]^{\alpha}|p(t)|dt+\sum^{\infty}[\tau(\theta_k)]^{\alpha}|q_k|<\infty, \tag{2.1}$$

then (1.1) *has a solution* $x(t)$ *satisfying*

$$\lim_{t\to\infty}\frac{x(t)}{t}=a\ne 0. \tag{2.2}$$

to Contents, page 1

Proof. Choose $t_1 \geq \max\{1, t_0\}$. In view of Lemma 1.3 by integrating (1.1) twice from t_0 to t, we obtain

$$x(t) = x(t_1) - x'(t_1)(t - t_1) - \sum_{t_1 \leq \theta_k < t} q_k |x(\tau(\theta_k))|^{\alpha-1} x((\tau(\theta_k)))(t - \theta_k) \\ - \int_{t_1}^{t} (t - s) p(s) |x(\tau(s))|^{\alpha-1} x((\tau(s))) ds, \quad t \geq t_1. \tag{2.3}$$

Set

$$u(t) = c + \sum_{t_1 \leq \theta_k < t} |q_k| |x(\tau(\theta_k))|^{\alpha} + \int_{t_1}^{t} |p(s)| |x(\tau(s))|^{\alpha} \, ds, \quad t \geq t_1, \tag{2.4}$$

where $c = |x(t_1)| + |x'(t_1)|$. Then

$$|x(t)| \leq t u(t), \quad t \geq t_1. \tag{2.5}$$

Let $t_2 \geq t_1$ be such that $\tau(t) \geq t_1$ for all $t \geq t_2$. Replacing t by $\tau(t)$ in (2.5) and using the increasing character of $u(t)$, we see that

$$|x(\tau(t))| \leq \tau(t) u(t), \quad t \geq t_2. \tag{2.6}$$

From (2.4), we also see that

$$u'(t) = |p(t)| |x(\tau(t))|^{\alpha}, \quad t \neq \theta_k, \tag{2.7}$$

$$\Delta u(t)|_{t=\theta_k} = |q_k| |x(\tau(\theta_k))|^{\alpha} \tag{2.8}$$

for $t \geq t_2$ and $\theta_k \geq t_2$. Now, in view of (2.6) and (2.8), an integration of (2.7) from t_2 to t leads to

$$u(t) \leq c + \int_{t_2}^{t} |p(s)| [\tau(s)]^{\alpha} [u(s)]^{\alpha} ds + \sum_{t_2 \leq \theta_k < t} |q_k| [\tau(\theta_k)]^{\alpha} [u(\theta_k)]^{\alpha}. \tag{2.9}$$

Applying Lemma 1.4 with

$$h(x) = x^{\alpha}, \qquad \lambda(s) = |p(s)| [\tau(s)]^{\alpha}, \qquad \lambda_k = |q_k| [\tau(\theta_k)]^{\alpha}, \tag{2.10}$$

we easily see that

$$u(t) \leq G^{-1} \left\{ G(c) + \int_{t_2}^{t} |p(s)| [\tau(s)]^{\alpha} ds + \sum_{t_2 \leq \theta_k < t} |q_k| [\tau(\theta_k)]^{\alpha} \right\}. \tag{2.11}$$

to Contents, page 1

Since

$$G(u) = \frac{u^{1-\alpha}}{1-\alpha} - \frac{u_0^{1-\alpha}}{1-\alpha}, \qquad G^{-1}(u) = \left[(1-\alpha)u + u_0^{1-\alpha}\right]^{1/(1-\alpha)}, \tag{2.12}$$

the inequality (2.11) becomes

$$u(t) \le \left[c^{1-\alpha} + (1-\alpha)\int_{t_1}^{t} |p(s)| [\tau(s)]^{\alpha} ds + (1-\alpha)\sum_{t_1 \le \theta_k < t} |q_k| [\tau(\theta_k)]^{\alpha}\right]^{1/(1-\alpha)}, \tag{2.13}$$

from which, on using (2.1), we have

$$u(t) \le c_1, \quad t \ge t_2, \tag{2.14}$$

where

$$c_1 = \left[c^{1-\alpha} + (1-\alpha)\int_{t_1}^{\infty} |p(s)| [\tau(s)]^{\alpha} ds + (1-\alpha)\sum_{t_1 \le \theta_k < \infty} |q_k| [\tau(\theta_k)]^{\alpha}\right]^{1/(1-\alpha)}. \tag{2.15}$$

In view of (2.5), (2.6), and (2.14) we see that

$$|x(t)| \le c_1 t, \quad |x(\tau(t))| \le c_1 \tau(t), \quad t \ge t_2. \tag{2.16}$$

To complete the proof it suffices to show that $x'(t)$ approaches a nonzero limit as t tends to ∞. To see this we integrate (1.1) from t_2 to t to get

$$x'(t) = x'(t_1) - \int_{t_2}^{t} p(s)|x(\tau(s))|^{\alpha-1} x(\tau(s)) ds - \sum_{t_2 \le \theta_k < t} q_k |x(\tau(\theta_k))|^{\alpha-1} x(\tau(\theta_k)). \tag{2.17}$$

Employing (2.16) we have

$$\begin{gathered} \int_{t_2}^{\infty} |p(s)x(\tau(s))|^{\alpha} ds \le c_1^{\alpha} \int_{t_2}^{\infty} |p(s)| [\tau(s)]^{\alpha} ds < \infty, \\ \sum_{t_2 \le \theta_k < \infty} |q_k x(\tau(\theta_k))|^{\alpha} \le c_1^{\alpha} \sum_{t_2 \le \theta_k < \infty} |q_k| [\tau(\theta_k)]^{\alpha} < \infty. \end{gathered} \tag{2.18}$$

Therefore, $\lim_{t\to\infty} x'(t) = L$ exists. Clearly, we can make $L \ne 0$ by requiring that

$$x'(t_2) > c_1^{\alpha}\left[\int_{t_2}^{\infty} |p(s)| [\tau(s)]^{\alpha} ds + \sum_{t_2 \le \theta_k < \infty} |q_k| [\tau(\theta_k)]^{\alpha}\right], \tag{2.19}$$

which is always possible by arranging t_2. □

to Contents, page 1

Theorem 2.2. *Suppose that p and $\{q_k\}$ are nonnegative. Then every solution of* (1.1) *is oscillatory if and only if*

$$\int^{\infty} [\tau(t)]^{\alpha} p(t)dt + \sum^{\infty} [\tau(\theta_k)]^{\alpha} q_k = \infty. \tag{2.20}$$

Proof. Let (2.20) fail to hold. Then, by Theorem 2.1 we see that there is a solution $x(t)$ which satisfies (2.2). Clearly, such a solution is nonoscillatory. This proves the necessity.

To show the sufficiency, suppose that (2.20) is valid but there is a nonoscillatory solution $x(t)$ of (1.1). We may assume that $x(t)$ is eventually positive; the case $x(t)$ being eventually negative is similar. Clearly, there exists $t_1 \geq t_0$ such that $x(\tau(t)) > 0$ for all $t \geq t_1$. From (1.1), we have that

$$x''(t) \leq 0 \quad \text{for } t \geq t_1, \ t \neq \theta_k. \tag{2.21}$$

Thus, $x'(t)$ is decreasing on every interval not containing $t = \theta_k$. From the impulse conditions in (1.1), we also have $\Delta x'(\theta_k) \leq 0$. Therefore, we deduce that $x'(t)$ is nondecreasing on $[t_1, \infty)$.

We may claim that $x'(t)$ is eventually positive. Because if $x'(t) < 0$ eventually, then $x(t)$ becomes negative for large values of t. This is a contradiction.

It is now easy to show that

$$x(t) \geq (t - t_1)x'(t), \quad t \geq t_1. \tag{2.22}$$

Therefore,

$$x(t) \geq \frac{t}{2}x'(t), \quad t \geq t_2 = 2t_1. \tag{2.23}$$

Let $t_3 \geq t_2$ be such that $\tau(t) \geq t_2$ for $t \geq t_3$. Using (2.23) and the nonincreasing character of $x'(t)$, we have

$$x(\tau(t)) \geq \frac{\tau(t)}{2}x'(t), \quad t \geq t_3, \tag{2.24}$$

and so, by (1.1),

$$x''(t) + 2^{-\alpha}p(t)[\tau(t)]^{\alpha}\left[x'(t)\right]^{\alpha} \leq 0, \quad t \neq \theta_k. \tag{2.25}$$

Dividing (2.25) by $[x'(t)]^{\alpha}$ and integrating from t_3 to t, we obtain

$$\begin{aligned}
&\sum_{t_3 \leq \theta_k < t} \left\{ \left[x'(\theta_k)\right]^{1-\alpha} - \left[x'(\theta_k) - q_k[x(\tau(\theta_k))]^{\alpha}\right]^{1-\alpha} \right\} \\
&\quad + \left[x'(t)\right]^{1-\alpha} - \left[x'(t_3)\right]^{1-\alpha} + (1-\alpha)2^{-\alpha}\int_{t_3}^{t} [\tau(t)]^{\alpha}p(s)ds \leq 0
\end{aligned} \tag{2.26}$$

to Contents, page 1

which clearly implies that

$$\sum_{t_3\le\theta_k<t} a_k + (1-\alpha)2^{-\alpha}\int_{t_3}^{t}[\tau(t)]^{\alpha}p(s)ds \le \left[x'(t_3)\right]^{1-\alpha}, \tag{2.27}$$

where

$$a_k = \left[x'(\theta_k)\right]^{1-\alpha}\left[1-\left(1-\frac{q_k[x(\tau(\theta_k))]^{\alpha}}{x'(\theta_k)}\right)\right]^{1-\alpha}. \tag{2.28}$$

Since $1-(1-u)^{1-\alpha} \ge (1-\alpha)u$ for $u \in (0,\infty)$ and $0<\alpha<1$, by taking

$$u = \frac{q_k[x(\tau(\theta_k))]^{\alpha}}{x'(\theta_k)}, \tag{2.29}$$

we see from (2.28) that

$$a_k \ge (1-\alpha)\frac{q_k[x(\tau(\theta_k))]^{\alpha}}{[x'(\theta_k)]^{\alpha}}. \tag{2.30}$$

But, (2.24) gives

$$x(\tau(\theta_k)) \ge \frac{\tau(\theta_k)}{2}x'(\tau(\theta_k)) \ge \frac{\tau(\theta_k)}{2}x'(\theta_k), \tag{2.31}$$

and hence

$$a_k \ge (1-\alpha)2^{-\alpha}[\tau(\theta_k)]^{\alpha}q_k. \tag{2.32}$$

Finally, (2.27) and (2.32) result in

$$\int_{t_3}^{\infty}[\tau(t)]^{\alpha}p(t)dt + \sum_{t_3<\theta_k<\infty}[\tau(\theta_k)]^{\alpha}q_k < \infty, \tag{2.33}$$

which contradicts (2.20). The proof is complete. □

Example 2.3. Consider the impulsive delay differential equation

$$\begin{gathered} x''(t) + (t-1)^{-2}|x(t-1)|^{-1/2}x(t-1) = 0, \quad t \ne k, \\ \Delta x'(t)\big|_{t=k} + (k-1)^{-1}|x(k-1)|^{-1/2}x(k-1) = 0, \\ \Delta x(t)|_{t=k} = 0, \end{gathered} \tag{2.34}$$

where $t \ge 2$ and $i \ge 2$.

to Contents, page 1

We see that $\tau(t) = t - 1$, $\alpha = 1/2$, $p(t) = (t-1)^{-2}$, and $q_k = (k-1)^{-1}$, $\theta_k = k$. Since

$$\int^{\infty} (t-1)^{-3/2} dt + \sum^{\infty} (k-1)^{-1/2} = \infty, \tag{2.35}$$

applying Theorem 2.2 we conclude that every solution of (2.34) is oscillatory.

We note that if the equation is not subject to any impulse condition, then, since

$$\int^{\infty} (t-1)^{-5/2} dt < \infty, \tag{2.36}$$

the equation

$$x''(t) + (t-1)^{-2} |x(t-1)|^{-1/2} x(t-1) = 0 \tag{2.37}$$

has a nonoscillatory solution by Theorem 1.1.

Let us now consider (1.1) when $\tau(t) = t$. That is,

$$\begin{gathered} x'' + p(t)|x|^{\alpha-1}x = 0, \quad t \neq \theta_k, \\ \Delta x'\big|_{t=\theta_k} + q_k |x|^{\alpha-1} x = 0, \\ \Delta x|_{t=\theta_k} = 0, \end{gathered} \tag{2.38}$$

where $0 < \alpha < 1$ and p q_k are given by (a) and (b).

The following theorem is an extension of Theorem 1.2. Note that no sign condition is imposed on $p(t)$ and $\{q_k\}$.

Theorem 2.4. *If*

$$\int^{\infty} t^{\beta} p(t) dt + \sum^{\infty} \theta_k^{\beta} q_k = \infty \tag{2.39}$$

for some $\beta \in [0, \alpha]$, *then every solution of* (2.38) *is oscillatory.*

Proof. Assume on the contrary that (2.38) has a nonoscillatory solution $x(t)$ such that $x(t) > 0$ for all $t \geq t_0$ for some $t_0 \geq 0$. The proof is similar when $x(t)$ is eventually negative. We set

$$w(t) = \left(t^{-1} x(t)\right)^{1-\alpha}, \quad t \geq t_0. \tag{2.40}$$

It is not difficult to see that

$$w'(t) = (\alpha - 1) t^{\alpha-2} [x(t)]^{1-\alpha} + (1-\alpha) t^{\alpha-1} [x(t)]^{-\alpha} x'(t), \quad t \neq \theta_k, \tag{2.41}$$

to Contents, page 1

and hence

$$\Delta w'\big|_{t=\theta_k} = (1-\alpha)q_k\theta_k^{\alpha-1}. \tag{2.42}$$

From (2.41), we have

$$\begin{aligned} t^{\beta-1-\alpha}\left(t^2 w'(t)\right)' &= (1-\alpha)t^{\beta}x''(t)x^{-\alpha}(t) \\ &\quad - \alpha(1-\alpha)t^{\beta-2}x^{-\alpha-1}(t)\left[tx'(t)-x(t)\right]^2, \end{aligned} \tag{2.43}$$

and so

$$t^{\beta-1-\alpha}\left(t^2 w'(t)\right)' \le (1-\alpha)t^{\beta}p(t), \quad t \ne \theta_k. \tag{2.44}$$

In view of (2.42), by a straightforward integration of (2.44), we have

$$\begin{aligned} \int_{t_0}^{t} s^{\beta-1-\alpha}\left(s^2 w'(s)\right)' ds &= s^{\beta-1-\alpha}s^2 w'(s)\Big|_{t_0}^{t} - \sum_{t_0\le\theta_k<t}\Delta\left(t^{\beta-\alpha+1}w'(t)\right)\Big|_{t=\theta_k} \\ &\quad - \int_{t_0}^{t}(\beta-1-\alpha)s^{\beta-\alpha}w'(s)ds \\ &= t^{\beta-\alpha+1}w'(t) - t_0^{\beta-\alpha+1}w'(t_0) - \sum_{t_0\le\theta_k<t}(1-\alpha)q_k\theta_k^{\beta} \\ &\quad -(\beta-\alpha-1)\left[s^{\beta-\alpha}w(s)\right]\Big|_{t_0}^{t} \\ &\quad + (\beta-\alpha)(\beta-\alpha-1)\int_{t_0}^{t} s^{\beta-1-\alpha}w(s)ds, \end{aligned} \tag{2.45}$$

which combined with (2.44) leads to

$$\begin{aligned} t^{\beta-\alpha+1}w'(t) &\le t_0^{\beta-\alpha+1}w'(t_0) - (\beta-\alpha+1)t_0^{\beta-\alpha}w(t_0) \\ &\quad + (1-\alpha)\left[\sum_{t_0\le\theta_k<t}\theta_k^{\beta}q_k + \int_{t_0}^{t} s^{\beta}p(s)ds\right]. \end{aligned} \tag{2.46}$$

Finally, by using (2.39) in the last inequality, we see that there is a $t_1 > t_0$ such that

$$w'(t) \le -t^{\alpha-\beta-1}, \quad t \ge t_1, \tag{2.47}$$

which, however, implies that $w(t) \to -\infty$ as $t \to \infty$, a contradiction with $x(t) > 0$. The proof is complete. □

to Contents, page 1

Example 2.5. Consider the impulsive differential equation

$$x'' + t^{-7/3}|x|^{-1/2}x = 0, \quad t \neq k,$$
$$\Delta x'\big|_{t=k} + k^{-1/6}|x|^{-1/2}x = 0, \tag{2.48}$$
$$\Delta x|_{t=k} = 0,$$

where $t \geq 1$ and $i \geq 1$.

We have that $p(t) = t^{7/3}$, $\alpha = 1/2$, and $q_k = k^{-1/6}$, $\theta_k = k$. Taking $\beta = 1/3$ we see from (2.38) that

$$\int^{\infty} t^{-2}dt + \sum^{\infty} k^{-1/3} = \infty. \tag{2.49}$$

Since the conditions of Theorem 2.4 are satisfied, every solution of (2.48) is oscillatory.

Note that if the impulses are absent, then, since

$$\int^{\infty} t^{-2}dt < \infty, \tag{2.50}$$

the equation

$$x'' + t^{-7/3}|x|^{-1/2}x = 0 \tag{2.51}$$

is oscillatory by Theorem 1.2.

Acknowledgment

This work was partially supported by METU-B AP (project no: 01-01-2011-003).

References

[1] T. Kusano and H. Onose, "Nonlinear oscillation of a sublinear delay equation of arbitrary order," *Proceedings of the American Mathematical Society*, vol. 40, pp. 219–224, 1973.
[2] H. E. Gollwitzer, "On nonlinear oscillations for a second order delay equation," *Journal of Mathematical Analysis and Applications*, vol. 26, pp. 385–389, 1969.
[3] V. N. Sevelo and O. N. Odaric, "Certain questions on the theory of the oscillation (non-oscillation) of the solutions of second order differential equations with retarded argument," *Ukrainskii Matematicheskii Zhurnal*, vol. 23, pp. 508–516, 1971 (Russian).
[4] S. Belohorec, "Two remarks on the properties of solutions of a nonlinear differential equation," *Acta Facultatis Rerum Naturalium Universitatis Comenianae/Mathematica*, vol. 22, pp. 19–26, 1969.
[5] D. D. Bainov, Yu. I. Domshlak, and P. S. Simeonov, "Sturmian comparison theory for impulsive differential inequalities and equations," *Archiv der Mathematik*, vol. 67, no. 1, pp. 35–49, 1996.
[6] K. Gopalsamy and B. G. Zhang, "On delay differential equations with impulses," *Journal of Mathematical Analysis and Applications*, vol. 139, no. 1, pp. 110–122, 1989.
[7] J. Yan, "Oscillation properties of a second-order impulsive delay differential equation," *Computers & Mathematics with Applications*, vol. 47, no. 2-3, pp. 253–258, 2004.

to Contents, page 1

[8] C. Yong-shao and F. Wei-zhen, "Oscillations of second order nonlinear ODE with impulses," *Journal of Mathematical Analysis and Applications*, vol. 210, no. 1, pp. 150–169, 1997.

[9] Z. He and W. Ge, "Oscillations of second-order nonlinear impulsive ordinary differential equations," *Journal of Computational and Applied Mathematics*, vol. 158, no. 2, pp. 397–406, 2003.

[10] C. Huang, "Oscillation and nonoscillation for second order linear impulsive differential equations," *Journal of Mathematical Analysis and Applications*, vol. 214, no. 2, pp. 378–394, 1997.

[11] J. Luo, "Second-order quasilinear oscillation with impulses," *Computers & Mathematics with Applications*, vol. 46, no. 2-3, pp. 279–291, 2003.

[12] A. Özbekler and A. Zafer, "Sturmian comparison theory for linear and half-linear impulsive differential equations," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 63, no. 5–7, pp. e289–e297, 2005.

[13] A. Özbekler and A. Zafer, "Picone's formula for linear non-selfadjoint impulsive differential equations," *Journal of Mathematical Analysis and Applications*, vol. 319, no. 2, pp. 410–423, 2006.

[14] G. Ballinger and X. Liu, "Permanence of population growth models with impulsive effects," *Mathematical and Computer Modelling*, vol. 26, no. 12, pp. 59–72, 1997.

[15] Z. Lu, X. Chi, and L. Chen, "Impulsive control strategies in biological control of pesticide," *Theoretical Population Biology*, vol. 64, no. 1, pp. 39–47, 2003.

[16] J. Sun, F. Qiao, and Q. Wu, "Impulsive control of a financial model," *Physics Letters A*, vol. 335, no. 4, pp. 282–288, 2005.

[17] S. Tang and L. Chen, "Global attractivity in a "food-limited" population model with impulsive effects," *Journal of Mathematical Analysis and Applications*, vol. 292, no. 1, pp. 211–221, 2004.

[18] S. Tang, Y. Xiao, and D. Clancy, "New modelling approach concerning integrated disease control and cost-effectivity," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 63, no. 3, pp. 439–471, 2005.

[19] V. Lakshmikantham, D. D. Baĭnov, and P. S. Simeonov, *Theory of Impulsive Differential Equations*, vol. 6 of *Series in Modern Applied Mathematics*, World Scientific, Teaneck, NJ, USA, 1989.

[20] A. M. Samoĭlenko and N. A. Perestyuk, *Impulsive Differential Equations*, vol. 14 of *World Scientific Series on Nonlinear Science. Series A: Monographs and Treatises*, World Scientific, River Edge, NJ, USA, 1995.

[21] M. Akhmetov and R. Sejilova, "The control of the boundary value problem for linear impulsive integro-differential systems," *Journal of Mathematical Analysis and Applications*, vol. 236, no. 2, pp. 312–326, 1999.

[22] D. Bainov and V. Covachev, *Impulsive Differential Equations with a Small Parameter*, vol. 24 of *Series on Advances in Mathematics for Applied Sciences*, World Scientific, River Edge, NJ, USA, 1994.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 635926, 12 pages
doi:10.1155/2011/635926

Research Article

Oscillatory Periodic Solutions for Two Differential-Difference Equations Arising in Applications

Rong Cheng

College of Mathematics and Physics, Nanjing University of Information Science and Technology, Nanjing 210044, China

Correspondence should be addressed to Rong Cheng, mathchr@163.com

Received 28 November 2010; Revised 31 January 2011; Accepted 2 March 2011

Academic Editor: Elena Braverman

We study the existence of oscillatory periodic solutions for two nonautonomous differential-difference equations which arise in a variety of applications with the following forms: $\dot{x}(t) = -f(t, x(t-r))$ and $\dot{x}(t) = -f(t, x(t-s)) - f(t, x(t-2s))$, where $f \in C(\mathbb{R} \times \mathbb{R}, \mathbb{R})$ is odd with respect to x, and $r, s > 0$ are two given constants. By using a symplectic transformation constructed by Cheng (2010) and a result in Hamiltonian systems, the existence of oscillatory periodic solutions of the above-mentioned equations is established.

1. Introduction and Statement of Main Results

Furumochi [1] studied the following equation:

$$\dot{x}(t) = a - \sin(x(t-r)), \tag{1.1}$$

with $t \geq 0, a \geq 0, r > 0$, which models phase-locked loop control of high-frequency generators and is widely applied in communication systems. Obviously, (1.1) is a special case of the following differential-difference equations:

$$\dot{x}(t) = -\alpha f(x(t-r)), \tag{1.2}$$

where α is a real parameter. In fact, a lot of differential-difference equations occurring widely in applications and describing many interesting types of phenomena can also be written in

to Contents, page 1

the form of (1.2) by making an appropriate change of variables. For example, the following differential-difference equation:

$$\dot{x}(t) = -\alpha x(t-1)(1+x(t)) \tag{1.3}$$

arises in several applications and has been studied by many researchers. Equation (1.3) was first considered by Cunningham [2] as a nonlinear growth model denoting a mathematical description of a fluctuating population. Subsequently, (1.3) was proposed by Wright [3] as occurring in the application of probability methods to the theory of asymptotic prime number density. Jones [4] states that (1.3) may also describe the operation of a control system working with potentially explosive chemical reactions, and quite similar equations arise in economic studies of business cycles. Moreover, (1.3) and its similar ones were studied in [5] on ecology.

For (1.3), we make the following change of variables:

$$y = \ln(1+x). \tag{1.4}$$

Then, (1.3) can be changed to the form of (1.2)

$$\dot{y}(t) = -f\left(y(t-1)\right), \tag{1.5}$$

where $f(y) = \alpha(e^y - 1)$.

Although (1.2) looks very simple on surface, Saupe's results [6] of a careful numerical study show that (1.2) displays very complex dynamical behaviour. Moreover, little of them has been proved to the best of the author's knowledge.

Due to a variety of applications, (1.2) attracts many authors to study it. In 1970s and 1980s of the last century, there has been a great deal of research on problems of the existence of periodic solutions [1, 4, 7–10], slowly oscillating solutions [11], stability of solutions [12–14], homoclinic solutions [15], and bifurcations of solutions [6, 16, 17] to (1.2).

Since, generally, the main tool used to conclude the existence of periodic solutions is various fixed-point theorems, here we want to mention Kaplan and Yorke's work on the existence of oscillatory periodic solutions of (1.5) in [7]. In [7], they considered the following equations:

$$\begin{gathered}\dot{x}(t) = -f(x(t-1)), \\ \dot{x}(t) = -f(x(t-1)) - f(x(t-2)),\end{gathered} \tag{1.6}$$

where f is continuous, $xf(x) > 0$ for $x \neq 0$, and f satisfies some asymptotically linear conditions at 0 and ∞. The authors introduced a new technique for establishing the existence of oscillatory periodic solutions of (1.6). They reduced the search for periodic solutions of (1.6) to the problem of finding periodic solutions for a related systems of ordinary differential equations. We will give more details about the reduction method in Section 2.

In 1990s of the last century and at the beginning of this century, some authors [18–21] applied Kaplan and Yorke's original ideas in [7] to study the existence and multiplicity of periodic solutions of (1.2) with more than two delays. See also [22, 23] for some other methods.

to Contents, page 1

The previous work mainly focuses on the autonomous differential-difference equation (1.2). However, some papers [13, 24] contain some interesting nonautonomous differential difference equations arising in economics and population biology where the delay r of (1.2) depends on time t instead of a positive constant. Motivated by the lack of more results on periodic solutions for nonautonomous differential-difference equations, in the present paper, we study the following equations:

$$\dot{x}(t) = -f(t, x(t-r)), \tag{1.7}$$

$$\dot{x}(t) = -f(t, x(t-s)) - f(t, x(t-2s)), \tag{1.8}$$

where $f(t,x) \in C(\mathbb{R} \times \mathbb{R}, \mathbb{R})$ is odd with respect to x and $r = \pi/2$, $s = \pi/3$. Here, we borrow the terminology "oscillatory periodic solution" for (1.7) and (1.8) since $f(t,x)$ is odd with respect to x.

Now, we state our main results as follows.

Theorem 1.1. *Suppose that $f(t,x) \in C(\mathbb{R}\times\mathbb{R}, \mathbb{R})$ is odd with respect to x and r-periodic with respect to t. Suppose that*

$$\lim_{x\to 0} \frac{f(t,x)}{x} = \omega_0(t), \qquad \lim_{x\to\infty} \frac{f(t,x)}{x} = \omega_\infty(t) \tag{1.9}$$

exist. Write $\alpha_0 = (1/r)\int_0^r \omega_0(t)dt$ and $\alpha_\infty = (1/r)\int_0^r \omega_\infty(t)dt$. Assume that

(H_1) $\alpha_0 \neq \pm k$, $\alpha_\infty \neq \pm k$, for all $k \in \mathbb{Z}^+$,

(H_2) there exists at least an integer k_0 with $k_0 \in \mathbb{Z}^+$ such that

$$\min\{\alpha_0, \alpha_\infty\} < \pm k_0 < \max\{\alpha_0, \alpha_\infty\}, \tag{1.10}$$

then (1.7) has at least one nontrivial oscillatory periodic solution x satisfying $x(t) = -x(t-\pi)$.

Theorem 1.2. *Suppose that $f(t,x) \in C(\mathbb{R}\times\mathbb{R}, \mathbb{R})$ is odd with respect to x and s-periodic with respect to t. Let $\omega_0(t)$ and $\omega_\infty(t)$ be the two functions defined in Theorem 1.1. Write $\beta_0 = (1/s)\int_0^s \omega_0(t)dt$ and $\beta_\infty = (1/s)\int_0^s \omega_\infty(t)dt$. Assume that*

(H_3) $\beta_0, 3\beta_0 \neq \pm k$, $\beta_\infty, 3\beta_\infty \neq \pm k$, for all $k \in \mathbb{Z}^+$,

(H_4) there exists at least an integer k_0 with $k_0 \in \mathbb{Z}^+$ such that

$$\min\{\beta_0, \beta_\infty\} < \pm k_0 < \max\{\beta_0, \beta_\infty\} \tag{1.11}$$

or

$$\min\{\beta_0, \beta_\infty\} < \pm\frac{k_0}{3} < \max\{\beta_0, \beta_\infty\}, \tag{1.12}$$

then (1.8) has at least one nontrivial oscillatory periodic solution x satisfying $x(t) = -x(t-\pi)$.

to Contents, page 1

Remark 1.3. Theorems 1.1 and 1.2 are concerned with the existence of periodic solutions for nonautonomous differential-difference equations (1.7) and (1.8). Therefore, our results generalize some results obtained in the references. We will use a symplectic transformation constructed in [25] and a theorem of [26] to prove our main results.

2. Proof of the Main Results

Consider the following nonautonomous Hamiltonian system:

$$\dot{z}(t) = J\nabla_z H(t,z), \tag{2.1}$$

where $J = \begin{pmatrix} 0 & -I_N \\ I_N & 0 \end{pmatrix}$ is the standard symplectic matrix, I_N is the identity matrix in $\mathbb{R}^N$, $\nabla_z H(t,z)$ denotes the gradient of $H(t,z)$ with respect to z, and $H \in C^1(\mathbb{R}\times\mathbb{R}^{2N},\mathbb{R})$ is the Hamiltonian function. Suppose that there exist two constant symmetric matrices h_0 and h_∞ such that

$$\begin{aligned} \nabla_z H(t,z) - h_0 z = o(|z|), \quad \text{as } |z| \longrightarrow 0, \\ \nabla_z H(t,z) - h_\infty z = o(|z|), \quad \text{as } |z| \longrightarrow \infty. \end{aligned} \tag{2.2}$$

We call the Hamiltonian system (2.1) asymptotically linear both at 0 and ∞ with constant coefficients h_0 and h_∞ because of (2.2).

Now, we show that the reduction method in [7] can be used to study oscillatory periodic solutions of (1.7) and (1.8). More precisely, let $x(t)$ be any solution of (1.7) satisfying $x(t) = -x(t-2r)$. Let $x_1(t) = x(t), x_2(t) = x(t-r)$, then $X(t) = (x_1(t), x_2(t))^\top$ satisfies

$$\frac{d}{dt}X(t) = A_2\Phi_1(t,X(t)), \quad \text{where } A_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{2.3}$$

and $\Phi_1(t,X) = (f(t,x_1), f(t,x_2))^\top$. What is more, if $X(t)$ is a solution of (2.3) with the following symmetric structure

$$x_1(t) = -x_2(t-r), \qquad x_2(t) = x_1(t-r), \tag{2.4}$$

then $x(t) = x_1(t)$ gives a solution to (1.7) with the property $x(t) = -x(t-2r)$. Thus, solving (1.7) within the class of the solutions with the symmetry $x(t) = -x(t-2r)$ is equivalent to finding solutions of (2.3) with the symmetric structure (2.4).

Since A_2 is indeed the standard symplectic matrix in the plane $\mathbb{R}^2$, the system (2.3) can be written as the following Hamiltonian system:

$$\dot{y}(t) = A_2\nabla_y H^*(t,y), \tag{2.5}$$

where $H^*(t,y) = \int_0^{y_1} f(t,x)dx + \int_0^{y_2} f(t,x)dx$ for each $y = (y_1,y_2)^\top \in \mathbb{R}^2$.

to Contents, page 1

From the assumptions of Theorem 1.1, we have

$$\begin{aligned} f(t,x) &= \omega_0(t)x + o(|x|) \quad \text{as } |x| \longrightarrow 0, \\ f(t,x) &= \omega_\infty(t)x + o(|x|) \quad \text{as } |x| \longrightarrow \infty. \end{aligned} \tag{2.6}$$

Hence, the gradient of the Hamiltonian function $H^*(t,y)$ satisfies

$$\begin{aligned} \nabla_y H^*(t,y) &= \omega_0(t)y + o(|y|) \quad \text{as } |y| \longrightarrow 0, \\ \nabla_y H^*(t,y) &= \omega_\infty(t)y + o(|y|) \quad \text{as } |y| \longrightarrow \infty. \end{aligned} \tag{2.7}$$

By (2.7), according to [25], there is a symplectic transformation $y = \Psi_1(t,z)$ under which the Hamiltonian system (2.5) can be transformed to the following Hamiltonian system:

$$\dot{z}(t) = A_2 \nabla_z \widetilde{H}(t,z), \tag{2.8}$$

satisfying

$$\begin{aligned} \nabla_z \widetilde{H}(t,z) &= \alpha_0 I_2 z + o(|z|) \quad \text{as } |z| \longrightarrow 0, \\ \nabla_z \widetilde{H}(t,z) &= \alpha_\infty I_2 z + o(|z|) \quad \text{as } |z| \longrightarrow \infty, \end{aligned} \tag{2.9}$$

where α_0 and α_∞ are two constants defined in Theorem 1.1.

By (2.9), we have the following.

Lemma 2.1. *The Hamiltonian system* (2.8) *is asymptotically linear both at 0 and ∞ with constant coefficients $\alpha_0 I_2$ and $\alpha_\infty I_2$.*

Let $x(t)$ be any solution of (1.8) satisfying $x(t) = -x(t-3s)$. Let $x_1(t) = x(t)$, $x_2(t) = x(t-s)$, and $x_3(t) = x(t-2s)$, then $Y(t) = (x_1(t), x_2(t), x_3(t))^\top$ satisfies

$$\frac{d}{dt}Y(t) = A_3 \Phi_2(t, Y(t)), \quad \text{where } A_3 = \begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix}, \tag{2.10}$$

and $\Phi_2(t,Y) = (f(t,x_1), f(t,x_2), f(t,x_3))^\top$.

Following the ideas in [18], (2.10) can be reduced to a two-dimensional Hamiltonian system

$$\dot{y}(t) = A_2 \nabla_y H^{**}(t,y), \tag{2.11}$$

where $H^{**}(t,y) = \int_0^{y_1} f(t,x)dx + \int_0^{y_2} f(t,x)dx + \int_0^{y_2-y_1} f(t,x)dx$ for each $y = (y_1, y_2)^\top \in \mathbb{R}^2$.

to Contents, page 1

From the assumptions of Theorem 1.1, (2.6), the gradient of the Hamiltonian function $H^{**}(t,y)$ satisfies

$$\begin{aligned}\nabla_y H^{**}(t,y) &= \omega_0(t)My + o(|y|) \quad \text{as } |y| \longrightarrow 0,\\ \nabla_y H^{**}(t,y) &= \omega_\infty(t)My + o(|y|) \quad \text{as } |y| \longrightarrow \infty,\end{aligned} \tag{2.12}$$

where $M = \left(\begin{smallmatrix} 2 & -1 \\ -1 & 2\end{smallmatrix}\right)$ is a symmetric positive definite matrix.

It follows from (2.12) and [25] that there exists a symplectic transformation $y = \Psi_2(t,z)$ under which the Hamiltonian system (2.11) can be changed to the following Hamiltonian system:

$$\dot{z}(t) = A_2 \nabla_z \widehat{H}(t,z), \tag{2.13}$$

satisfying

$$\begin{aligned}\nabla_z \widehat{H}(t,z) &= \beta_0 Mz + o(|z|) \quad \text{as } |z| \longrightarrow 0,\\ \nabla_z \widehat{H}(t,z) &= \beta_\infty Mz + o(|z|) \quad \text{as } |z| \longrightarrow \infty,\end{aligned} \tag{2.14}$$

where β_0 and β_∞ are two constants defined in Theorem 1.2.

Then, (2.14) yields the following.

Lemma 2.2. *The Hamiltonian system* (2.13) *is asymptotically linear both at 0 and ∞ with constant coefficients $\beta_0 M$ and $\beta_\infty M$.*

Remark 2.3. In order to find periodic solutions of (1.7) and (1.8), we only need to seek periodic solutions of the Hamiltonian systems (2.8) and (2.13) with the symmetric structure (2.4), respectively.

In the rest of this paper, we will work in the Hilbert space $E = W^{1/2,2}(S^1, \mathbb{R}^2)$, which consists of all $z(t)$ in $L^2(S^1, \mathbb{R}^2)$ whose Fourier series

$$z(t) = a_0 + \sum_{k=1}^{+\infty} (a_k \cos kt + b_k \sin kt) \tag{2.15}$$

satisfies

$$|a_0|^2 + \frac{1}{2}\sum_{k=1}^{+\infty} k\left(|a_k|^2 + |b_k|^2\right) < +\infty. \tag{2.16}$$

The inner product on E is defined by

$$\langle z_1, z_2 \rangle = \left(a_0^{(1)}, a_0^{(2)}\right) + \frac{1}{2}\sum_{k=1}^{\infty} k\left[\left(a_k^{(1)}, a_k^{(2)}\right) + \left(b_k^{(1)}, b_k^{(2)}\right)\right], \tag{2.17}$$

to Contents, page 1

where $z_i = a_0^{(i)} + \sum_{k=1}^{+\infty}(a_k^{(i)}\cos kt + b_k^{(i)}\sin kt)$ $(i = 1, 2)$, the norm $\|z\|^2 = \langle z, z\rangle$, and $(\cdot,\cdot)$ denotes the inner product in $\mathbb{R}^2$.

In order to obtain solutions of (2.8) with the symmetric structure (2.4), we define a matrix T_2 with the following form:

$$T_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{2.18}$$

Then, by T_2, for any $z(t) \in E$, define an action δ_1 on z by

$$\delta_1 z(t) = T_2 z(t - r). \tag{2.19}$$

Then by a direct computation, we have that $\delta_1^2 z(t) = -z(t - 2r) = -z(t - \pi)$, $\delta_1^4 z(t) = z(t)$, and $G = \{\delta_1, \delta_1^2, \delta_1^3, \delta_1^4\}$ is a compact group action over E. If $\delta_1 z(t) = z(t)$ holds, then through a straightforward check, we have that $z(t)$ has the symmetric structure (2.4).

Lemma 2.4. *Write $SE = \{z \in E : \delta_1 z(t) = z(t)\}$, then SE is a subspace of E with the following form:*

$$\begin{aligned} SE = \Bigg\{ z(t) = \sum_{k=1}^{\infty}&\left(a_{2k-1}\cos(2k-1)t + b_{2k-1}\sin(2k-1)t\right) : \\ &a_{2k-1,1} = (-1)^{k+1} b_{2k-1,2},\, b_{2k-1,1} = (-1)^k a_{2k-1,2} \Bigg\}, \end{aligned} \tag{2.20}$$

where $a_{2k-1} = (a_{2k-1,1}, a_{2k-1,2})^\top$ and $b_{2k-1} = (b_{2k-1,1}, b_{2k-1,2})^\top$.

Proof. Write $z(t) = (z_1(t), z_2(t))^\top$, where $z_1(t) = a_{0,1} + \sum_{k=1}^{+\infty}(a_{k,1}\cos kt + b_{k,1}\sin kt)$, $z_2(t) = a_{0,2} + \sum_{k=1}^{+\infty}(a_{k,2}\cos kt + b_{k,2}\sin kt)$. By $\delta_1 z = z$ and the definition of the action δ_1, we have

$$(z_1(t), z_2(t))^\top = \left(-z_2\left(t - \frac{\pi}{2}\right), z_1\left(t - \frac{\pi}{2}\right)\right)^\top, \tag{2.21}$$

which yields

$$\begin{aligned} &a_{0,1} + \sum_{k=1}^{+\infty}(a_{k,1}\cos kt + b_{k,1}\sin kt) \\ &= \begin{cases} -a_{0,2} - \displaystyle\sum_{n=1}^{+\infty}(-1)^n\left[a_{2n,2}\cos 2nt + b_{2n,2}\sin 2nt\right], & \text{for } k = 2n \text{ is even}, \\ -a_{0,2} - \displaystyle\sum_{n=1}^{+\infty}(-1)^{n-1}\left[a_{2n-1,2}\sin(2n-1)t - b_{2n-1,2}\cos(2n-1)t\right], & \text{for } k = 2n-1 \text{ is odd}. \end{cases} \end{aligned} \tag{2.22}$$

to Contents, page 1

Then, we have

$$\begin{gathered} a_{0,1} = -a_{0,2}, \qquad a_{2n,1} = (-1)^{n+1} a_{2n,2}, \qquad b_{2n,1} = (-1)^{n+1} b_{2n,2}, \\ a_{2n-1,1} = (-1)^{n+1} b_{2n-1,2}, \qquad b_{2n-1,1} = (-1)^{n} a_{2n-1,2}. \end{gathered} \tag{2.23}$$

Similarly, by $z_2(t) = z_1(t - (\pi/2))$, one has

$$\begin{gathered} a_{0,2} = a_{0,1}, \qquad a_{2n,2} = (-1)^{n} a_{2n,1}, \qquad b_{2n,2} = (-1)^{n} b_{2n,1}, \\ a_{2n-1,2} = (-1)^{n} b_{2n-1,1}, \qquad b_{2n-1,2} = (-1)^{n-1} a_{2n-1,1}. \end{gathered} \tag{2.24}$$

Therefore, $a_{0,2} = a_{0,1} = 0$, $a_{2n,1} = (-1)^{n+1} a_{2n,2} = (-1)^{n+1}(-1)^{n} a_{2n,1}$, that is, $a_{2n,1} = 0$. Similarly, $a_{2n,2} = b_{2n,1} = b_{2n,2} = 0$. Thus, for $z(t) \in \mathrm{SE}$,

$$z(t) = \sum_{k=1}^{\infty} \left[a_{2k-1} \cos(2k-1)t + b_{2k-1} \sin(2k-1)t\right], \tag{2.25}$$

where $a_{2k-1,1} = (-1)^{k+1} b_{2k-1,2}$, $b_{2k-1,1} = (-1)^{k} a_{2k-1,2}$.

Moreover, for any $z_1(t), z_2(t) \in \mathrm{SE}$,

$$\begin{aligned} \delta_1(z_1 + z_2) &= T_2(z_1(t-r) + z_2(t-r)) \\ &= T_2(z_1(t-r)) + T_2(z_2(t-r)) \\ &= \delta_1 z_1 + \delta_1 z_2. \end{aligned} \tag{2.26}$$

And for any $c \in \mathbb{R}$, $\delta_1(cz(t)) = T_2 cz(t-r) = cT_2 z(t-r) = c\delta_1 z(t)$. Thus, SE is a subspace of E. This completes the proof of Lemma 2.4. □

For the Hamiltonian system (2.13), we define another action matrix T_2^* with the following form:

$$T_2^* = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}. \tag{2.27}$$

Then, by T_2^*, for any $z(t) \in E$, define an action δ_2 on z by

$$\delta_2 z(t) = T_2^* z(t-s). \tag{2.28}$$

Then, by a direct computation, we have that $\delta_2^3 z(t) = -z(t-3s) = -z(t-\pi)$, $\delta_2^6 z(t) = z(t)$ and $G = \{\delta_2, \delta_2^2, \delta_2^3, \delta_2^4, \delta_2^5, \delta_2^6\}$ is a compact group action over E. If $\delta_2 z(t) = z(t)$ holds, then through a direct check, we have that $z(t)$ has the symmetric structure (2.4).

to Contents, page 1

Remark 2.5. By $\delta_2^3 z(t) = -z(t-3s) = -z(t-\pi)$ and the definition of δ_2, the set $\{z \in E : \delta_2 z(t) = z(t)\}$ has the same structure (2.20), where the relation between the Fourier coefficients of the first component z_1 and the second component z_2 is slightly different with the elements in $\{z \in E : \delta_1 z(t) = z(t)\}$. We denote it also by SE which is a subspace of E.

Denote by $M^-(h)$, $M^+(h)$, and $M^0(h)$ the number of the negative, the positive, and the zero eigenvalues of a symmetric matrix h, respectively. For a constant symmetric matrix h, we define our index as

$$i^-(h) = \sum_{k=1}^{\infty} \left(M^-(T_k(h) - 2)\right),$$
$$i^0(h) = \sum_{k=1}^{\infty} M^0(T_k(h)), \tag{2.29}$$

where

$$T_k(h) = \begin{pmatrix} -h & -kJ \\ kJ & -h \end{pmatrix}. \tag{2.30}$$

Observe that for k large enough, $M^-(T_k(h)) = 2$ and $M^0(T_k(h)) = 0$. In fact, write

$$T_k(h) = \begin{pmatrix} -h & -kJ \\ kJ & -h \end{pmatrix} = k\begin{pmatrix} 0 & J^\top \\ J & 0 \end{pmatrix} - \begin{pmatrix} h & 0 \\ 0 & h \end{pmatrix}. \tag{2.31}$$

Notice that $-J = J^\top$. If $k > 0$ is sufficiently large, then $M^- = M^+ = 2$, which are the indices of the first matrix in (2.31). Furthermore, if k decreases, these indices can change only at those values of k, for which the matrix $T_k(h)$ is singular, that is, $M^0(T_k(h)) \neq 0$. This happens exactly for those values of $k \in \mathbb{R}$ for which ik is a pure imaginary eigenvalue of Jh. Indeed assume $(\xi_1, \xi_2) \in \mathbb{R}^2 \times \mathbb{R}^2$ is an eigenvector of $T_k(h)$ with eigenvalue 0, then by $J^\top = -J$, one has $h\xi_1 + kJ\xi_2 = 0$ and $h\xi_2 - kJ\xi_1 = 0$. Thus, $h(\xi_1 + i\xi_2) = kJ(i\xi_1 - \xi_2) = ikJ(\xi_1 + i\xi_2)$; therefore, $Jh(\xi_1 + i\xi_2) = -ik(\xi_1 + i\xi_2)$. Therefore, $\pm ik \in \sigma(Jh)$, as claimed. Hence, $i^-(h)$ and $i^0(h)$ are well defined.

The following theorem of [26] on the existence of periodic solutions for the Hamiltonian system (2.1) will be used in our discussion.

Theorem A. *Let $H \in C^1(\mathbb{R} \times \mathbb{R}^{2N}, \mathbb{R})$ be 2π-periodic in t and satisfy* (2.2)*. If $i^0(h_0) = i^0(h_\infty) = 0$ and $i^-(h_0) \neq i^-(h_\infty)$, then the Hamiltonian system* (2.1) *has at least one nontrivial periodic solution.*

Now, we claim the following.

Lemma 2.6. *If z is a solution of the Hamiltonian system* (2.8) ((2.13)) *in SE, then $y = \Psi_1(t, z)$ ($y = \Psi_2(t, z)$) is the solution of the Hamiltonian system* (2.5) ((2.11)) *with the symmetric structure* (2.4)*, respectively.*

Proof. By Lemma 2.4, any $z \in \mathrm{SE}$ has the structure (2.4). We only need to show $\delta_1 y = y$ or $\delta_2 y = y$, that is, $T_2\Psi_1(t,z) = \Psi_1(t,T_2 z)$ or $T_2^*\Psi_2(t,z) = \Psi_2(t,T_2^* z)$, which can be verified directly by the constructions of the symplectic transformations $\Psi_1(t,z)$ and $\Psi_2(t,z)$, respectively. Please see [25] for details. □

We denote the matrix αI_2 by α for convenience. We prove the following lemma.

Lemma 2.7. (1) *Suppose that* (H_1) *and* (H_3) *hold, then* $i^0(\alpha_0) = i^0(\alpha_\infty) = i^0(\beta_0 M) = i^0(\beta_\infty M) = 0$.

(2) *Suppose that* (H_1) *and* (H_2) *hold, then* $i^-(\alpha_0) \neq i^-(\alpha_\infty)$.

(3) *Suppose that* (H_3) *and* (H_4) *hold, then* $i^-(\beta_0 M) \neq i^-(\beta_\infty M)$.

Proof. For any $\alpha, \beta \in \mathbb{R}$, let $\sigma(T_k(\alpha))$ and $\sigma(T_k(\beta M))$ denote the spectra of $T_k(\alpha)$ and $T_k(\beta M)$, respectively. Denote by λ and γ the elements of $\sigma(T_k(\alpha))$ and $\sigma(T_k(\beta M))$, respectively, then

$$\begin{aligned}
\det(\lambda I_4 - T_k(\alpha)) &= \det\Big((\lambda+\alpha)^2 I_2 - k^2 I_2\Big) \\
&= \det((\lambda+\alpha)I_2 - kI_2)\det((\lambda+\alpha)I_2 + kI_2), \\
\det\big(\gamma I_4 - T_k(\beta M)\big) &= \det\Big(\big(\gamma I_2 + \beta M\big)^2 - k^2 I_2\Big) \\
&= \det\big((\gamma I_2 + \beta M) - kI_2\big)\det\big((\gamma I_2 + \beta M) + kI_2\big) \\
&= \det\Big((\gamma + 2\beta - k)^2 - \beta^2\Big)\det\Big((\gamma + 2\beta + k)^2 - \beta^2\Big).
\end{aligned} \tag{2.32}$$

The above computation of determinant shows that

$$\sigma(T_k(\alpha)) = \{\lambda = \pm k - \alpha : k \in \mathbb{Z}^+\}, \tag{2.33}$$

$$\sigma\big(T_k(\beta M)\big) = \big\{\gamma = \pm k - \beta, \pm k - 3\beta : k \in \mathbb{Z}^+\big\}. \tag{2.34}$$

Case 1. From (2.33), if $\alpha_0 \neq \pm k$, for all $k \in \mathbb{Z}^+$, then $\lambda \neq 0$, where λ is the eigenvalue of $T_k(\alpha_0)$. That means $M^0(T_k(\alpha_0)) = 0$ for $k \geq 1$. Thus, $i^0(\alpha_0) = \sum_{k=1}^{\infty} M^0(T_k(\alpha_0)) = 0$. Similarly, we have $i^0(\alpha_\infty) = i^0(\beta_0 M) = i^0(\beta_\infty M) = 0$.

Case 2. Without loss of generality, we suppose that $\alpha_0 < \alpha_\infty$. By the conditions (H_1) and (H_2),

$$\alpha_0 < k_0 < \alpha_\infty. \tag{2.35}$$

Since $\alpha_0 < k_0$, by (2.33), $M^-(T_{k_0}(\alpha_0)) \leq 2$. By $-k_0 < k_0 < \alpha_\infty$ and (2.33), $M^-(T_{k_0}(\alpha_\infty)) = 4$, that is,

$$M^-(T_{k_0}(\alpha_0)) + 2 \leq M^-(T_{k_0}(\alpha_\infty)). \tag{2.36}$$

For each $k \neq k_0$ and from (2.33), one can check easily that $M^-(T_k(\alpha_0)) \leq M^-(T_k(\alpha_\infty))$. Hence, one has $\sum_{k=1}^{\infty}(M^-(T_k(\alpha_0)) - 2) < \sum_{k=1}^{\infty}(M^-(T_k(\alpha_\infty)) - 2)$, since $M^-(T_k(\alpha)) = 2$ for k large enough. This yields that $i^-(\alpha_0) < i^-(\alpha_\infty)$. Then, property (2) holds.

to Contents, page 1

Case 3. By the conditions *(H$_3$)* and *(H$_4$)*, without loss of generality, we suppose that $\beta_0 < \beta_\infty$ and

$$\beta_0 < k_0 < \beta_\infty. \tag{2.37}$$

Since $\beta_0 < k_0$, by (2.34), $M^-(T_{k_0}(\beta_0 M)) \le 3$. By $-k_0 < k_0 < \beta_\infty < 3\beta_\infty$ and (2.34), one has $M^-(T_{k_0}(\beta_\infty M)) = 4$, that is,

$$M^-\left(T_{k_0}\left(\beta_0 M\right)\right) + 1 \le M^-\left(T_{k_0}\left(\beta_\infty M\right)\right). \tag{2.38}$$

For each $k \neq k_0$ and from (2.34), it is easy to see that $k - \beta_\infty < k - \beta_0$ and $k - 3\beta_\infty < k - 3\beta_0$. Then, by the definition of $M^-(T_k(\beta M))$, we have $M^-(T_k(\beta_0 M)) \le M^-(T_k(\beta_\infty M))$. Therefore, we have

$$\sum_{k=1}^{\infty}\left(M^-\left(T_k\left(\beta_0 M\right)\right) - 2\right) < \sum_{k=1}^{\infty}\left(M^-\left(T_k\left(\beta_\infty M\right)\right) - 2\right), \tag{2.39}$$

since $M^-(T_k(\beta M)) = 2$ for k large enough. This implies that $i^-(\beta_0 M) < i^-(\beta_\infty M)$. Then, property (3) holds. □

Now, we are ready to prove the main results. We first give the proof of Theorem 1.1.

Proof of Theorem 1.1. Solutions of (2.8) in SE are indeed nonconstant classic 2π-periodic solutions with the symmetric structure (2.4), and hence they give solutions of (1.7) with the property $x(t - \pi) = -x(t)$. Therefore, we will seek solutions of (2.8) in SE.

Now, Theorem 1.1 follows from Lemmas 2.1, 2.6, and 2.7 and Theorem A. □

Proof of Theorem 1.2. Obviously, Theorem 1.2 follows from Lemmas 2.2, 2.6, and 2.7 and Theorem A. □

Acknowledgments

The author thanks the referee for carefully reading of the paper and giving valuable suggestions. This work is supported by the National Natural Science Foundation of China (11026212). This paper was typeset using AMS-LATEX.

References

[1] T. Furumochi, "Existence of periodic solutions of one-dimensional differential-delay equations," *The Tôhoku Mathematical Journal*, vol. 30, no. 1, pp. 13–35, 1978.
[2] W. J. Cunningham, "A nonlinear differential-difference equation of growth," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 40, no. 8, pp. 708–713, 1954.
[3] E. M. Wright, "A non-linear difference-differential equation," *Journal für die Reine und Angewandte Mathematik*, vol. 194, pp. 66–87, 1955.
[4] G. S. Jones, "The existence of periodic solutions of $f'(x) = -\alpha f(x(t-1))\{1 + f(x)\}$," *Journal of Mathematical Analysis and Applications*, vol. 5, pp. 435–450, 1962.
[5] R. May, *Stablity and Complexity in Model Ecosystems*, Princeton University Press, Princeton, NJ, USA, 1973.

to Contents, page 1

[6] D. Saupe, "Global bifurcation of periodic solutions of some autonomous differential delay equation," in *Forschungsschwerpunkt Dynamische Systems*, Report Nr. 71, University of Bremen, 1982.
[7] J. L. Kaplan and J. A. Yorke, "Ordinary differential equations which yield periodic solutions of differential delay equations," *Journal of Mathematical Analysis and Applications*, vol. 48, pp. 317–324, 1974.
[8] R. D. Nussbaum, "Periodic solutions of special differential equations: an example in nonlinear functional analysis," *Proceedings of the Royal Society of Edinburgh A*, vol. 81, no. 1-2, pp. 131–151, 1978.
[9] R. D. Nussbaum, "Uniqueness and nonuniqueness for periodic solutions of $x'(t) = -g(x(t-1))$," *Journal of Differential Equations*, vol. 34, no. 1, pp. 25–54, 1979.
[10] S. Chapin, "Periodic solutions of differential-delay equations with more than one delay," *The Rocky Mountain Journal of Mathematics*, vol. 17, no. 3, pp. 555–572, 1987.
[11] H.-O. Walther, "Density of slowly oscillating solutions of $\dot{x}(t) = -f(x(t-1))$," *Journal of Mathematical Analysis and Applications*, vol. 79, no. 1, pp. 127–140, 1981.
[12] J. L. Kaplan and J. A. Yorke, "On the stability of a periodic solution of a differential delay equation," *SIAM Journal on Mathematical Analysis*, vol. 6, pp. 268–282, 1975.
[13] J. A. Yorke, "Asymptotic stability for one dimensional differential-delay equations," *Journal of Differential Equations*, vol. 7, pp. 189–202, 1970.
[14] P. Dormayer, "The stability of special symmetric solutions of $\dot{x}(t) = \alpha f(x(t-1))$ with small amplitudes," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 14, no. 8, pp. 701–715, 1990.
[15] H.-O. Walther, "Homoclinic solution and chaos in $\dot{x}(t) = f(x(t-1))$," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 5, no. 7, pp. 775–788, 1981.
[16] R. D. Nussbaum, "A Hopf global bifurcation theorem for retarded functional differential equations," *Transactions of the American Mathematical Society*, vol. 238, pp. 139–164, 1978.
[17] M. Han, "Bifurcations of periodic solutions of delay differential equations," *Journal of Differential Equations*, vol. 189, no. 2, pp. 396–411, 2003.
[18] J. Li, X.-Z. He, and Z. Liu, "Hamiltonian symmetric groups and multiple periodic solutions of differential delay equations," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 35, no. 4, pp. 457–474, 1999.
[19] J. Li and X.-Z. He, "Multiple periodic solutions of differential delay equations created by asymptotically linear Hamiltonian systems," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 31, no. 1-2, pp. 45–54, 1998.
[20] G. Fei, "Multiple periodic solutions of differential delay equations via Hamiltonian systems (I)," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 65, no. 1, pp. 25–39, 2006.
[21] G. Fei, "Multiple periodic solutions of differential delay equations via Hamiltonian systems (II)," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 65, no. 1, pp. 40–58, 2006.
[22] J. Llibre and A.-A. Tarţa, "Periodic solutions of delay equations with three delays via bi-Hamiltonian systems," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 64, no. 11, pp. 2433–2441, 2006.
[23] S. Jekel and C. Johnston, "A Hamiltonian with periodic orbits having several delays," *Journal of Differential Equations*, vol. 222, no. 2, pp. 425–438, 2006.
[24] J. Bélair and M. C. Mackey, "Consumer memory and price fluctuations in commodity markets: an integrodifferential model," *Journal of Dynamics and Differential Equations*, vol. 1, no. 3, pp. 299–325, 1989.
[25] R. Cheng, "Symplectic transformations and a reduction method for asymptotically linear Hamiltonian systems," *Acta Applicandae Mathematicae*, vol. 110, no. 1, pp. 209–214, 2010.
[26] S. J. Li and J. Q. Liu, "Morse theory and asymptotic linear Hamiltonian system," *Journal of Differential Equations*, vol. 78, no. 1, pp. 53–73, 1989.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 928587, 12 pages
doi:10.1155/2011/928587

Research Article

Periodic Problems of Difference Equations and Ergodic Theory

B. A. Biletskyi,[1] A. A. Boichuk,[2] and A. A. Pokutnyi[2]

[1] *Department of Inductive Modeling and Control Glushkov, Institute of Cybernetics NAS Ukraine, No. 40 Glushkov Avenue, Kyiv 03680, Ukraine*
[2] *Laboratory of boundary value problems of differential equations theory, Institute of Mathematics of NAS of Ukraine, No. 3 Tereshenkivska Street, Kyiv 01601, Ukraine*

Correspondence should be addressed to A. A. Boichuk, boichuk@imath.kiev.ua

Received 1 February 2011; Accepted 11 May 2011

Academic Editor: Miroslava Růžičková

The necessary and sufficient conditions for solvability of the family of difference equations with periodic boundary condition were obtained using the notion of relative spectrum of linear bounded operator in the Banach space and the ergodic theorem. It is shown that when the condition of existence is satisfied, then such periodic solutions are built using the formula for the generalized inverse operator to the linear limited one.

1. The Problem and The Main Statement

The problem of existence of periodic solutions for classes of equations is well known. Though it is hard to mention all the contributors in a single paper, we would like to mark out well-developed Floke theory [1], which is used in analysis of linear differential equation systems by the means of monodromy matrix. Operator analogy of such theory in noncritical case (when there is single solution) for differential equations in Banach space was developed by Daletskyi and Krein [2].

This paper is dedicated to obtaining analogous conditions for a family of difference equations in Banach space and to building representations of corresponding solutions. The proposed approach allows obtaining solutions for both critical and noncritical cases. Note that this problem can be approached using well-developed pseudoinverse techniques in theory of boundary value problems [3]. In this paper we firstly build a new representation of the pseudoinverse operator based on results of ergodic theory, and then we provide the necessary and sufficient conditions that guarantee the existence of the corresponding solutions.

to Contents, page 1

Let $\mathbf{B}$-complex Banach space with norm $\|\cdot\|$ and zero-element $\overline{0}$; $\mathcal{L}(\mathbf{B})$-Banach space of bounded linear operators from $\mathbf{B}$ to $\mathbf{B}$. In this paper we consider existence of periodic solutions of the equation

$$x_{n+1} = \lambda A_{n+1} x_n + h_{n+1}, \quad n \geqslant 0, \tag{1.1}$$

with periodicity condition

$$x_0 = x_m, \tag{1.2}$$

where $A_n \in \mathcal{L}(\mathbf{B})$, $A_{n+m} = A_n$, for all $n \geqslant 0$, λ is a complex parameter, and $\{h_n\}_{n=0}^{\infty}$ is a sequence in $\mathbf{B}$. The solution of the corresponding homogeneous equation to (1.1) has the following form [4]:

$$x_m(\lambda) = \Phi(m, n, \lambda) x_n(\lambda), \quad m \geqslant n, \tag{1.3}$$

where

$$\Phi(m, n, \lambda) = \lambda^{m-n} A_{m+1} A_m \cdots A_{n+1}, \quad m > n \tag{1.4}$$

is evolution operator for problem (1.1); $\Phi(m, m, \lambda) = E$, where E is identity operator. Let us remark that $U(m, \lambda) = \Phi(m, 0, \lambda)$, $U(0, \lambda) = E$ and $U(k+n, \lambda) = U(k, \lambda) U(n, \lambda)$. Operator $U(m, \lambda)$ is traditionally called monodromy operator.

We can represent [4] the solution (1.1) with arbitrary initial condition $x(0, \lambda) = x_0$, $x_0 \in \mathbf{B}$ in the form

$$x_k(\lambda) = \Phi(k, 0, \lambda) x_0 + g(k, \lambda), \tag{1.5}$$

where

$$g(k, \lambda) = \sum_{i=0}^{k} \Phi(k, i, \lambda) h_i. \tag{1.6}$$

If we substitute this representation in boundary condition (1.2), we obtain operator equation

$$x_0(\lambda) - x_m(\lambda) = x_0 - \Phi(m, 0, \lambda) x_0 - g(m, \lambda) = \overline{0}. \tag{1.7}$$

According to notations, we get operator equation

$$(E - U(m, \lambda)) x_0 = g(m, \lambda). \tag{1.8}$$

Boundary value problem (1.1), (1.2) has periodic solution if and only if operator equation (1.8) is solvable.

to Contents, page 1

Following the paper [5], point λ is called *right stable point* if monodromy operator satisfies inequality $\{\|U^n(m,\lambda)\| \leqslant c,\ n \geqslant 0\}$.

Denote $\rho_{NS}(E - U(m,\lambda)) = \{\lambda \in \mathbb{C} : R(E - U(m,\lambda)) = \overline{R(E - U(m,\lambda))}\}$ (this set coincides with the set of all $\lambda \in \mathbb{C}$ such that operator $E - U(m,\lambda)$ is normally solvable). It follows easily that resolvent set $\rho(E - U(m,\lambda))$ of the operator $E - U(m,\lambda)$ lies in $\rho_{NS}(E - U(m,\lambda))$.

In the sequel we assume that **B** is reflexive for simplicity [6].

The main result of this paper is contained in Theorem 1.1.

Theorem 1.1. *Let $\lambda \in \rho_{NS}(E - U(m,\lambda))$ be right stability point for* (1.1). *Then*

(a) *boundary value problem* (1.1), (1.2) *has solutions if and only if sequence $\{h_n\}_{n\in\mathbb{Z}_+}$, $h_n \in \mathbf{B}$ satisfies condition*

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n}\sum_{i=0}^{m} U^k(m,\lambda)\Phi(m,i,\lambda)h_i}{n} = 0, \tag{1.9}$$

(b) *under condition* (1.9), *solutions of boundary value problem* (1.1), (1.2) *have the following form:*

$$x_n = U(n,\lambda)\lim_{k\to\infty}\frac{\sum_{m=1}^{k} U^m(k,\lambda)}{k}c + U(n,\lambda)G(n,\lambda)[h_n], \tag{1.10}$$

*where c is an arbitrary element of Banach space **B**, $G(n,\lambda)$-generalized Green operator of boundary value* (1.1), (1.2), *which is defined by equality*

$$\begin{aligned} G(n,\lambda)[h_n] = {} & \sum_{k=0}^{\infty}(1-\mu)^k\left\{\sum_{l=0}^{\infty}\mu^{-l-1}(U(m,\lambda) - U_0(\lambda))^l\right\}^{k+1}\sum_{i=0}^{m}\Phi(m,i,\lambda)h_i \\ & - U_0(\lambda)\sum_{i=0}^{m}\Phi(m,i,\lambda)h_i + \sum_{i=0}^{n}\Phi(n,i,\lambda)h_i. \end{aligned} \tag{1.11}$$

2. Auxiliary Result

Let us formulate and prove a number of auxiliary lemmas, which entail the theorem.

Lemma 2.1. *If $\lambda \in \rho_{NS}(E - U(m,\lambda))$, then boundary value problem* (1.1), (1.2) *is solvable if and only if sequence h_n satisfies the condition*

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n}\sum_{i=0}^{m} U^k(m,\lambda)\Phi(m,i,\lambda)h_i}{n} = 0. \tag{2.1}$$

to Contents, page 1

Proof. From the assumption above it follows that the conditions of statistical ergodic theorem hold [6]. Then

$$R(E-U(m,\lambda))=\left\{x\in\mathbf{B}:\lim_{n\to\infty}U_n(m,\lambda)x=\overline{0},U_n(m,\lambda)=\frac{\sum_{k=1}^{n}U^k(m,\lambda)}{n}\right\}. \tag{2.2}$$

It follows from the equation above that element $g(m,\lambda)$ lies in value set of the operator $E-U(m,\lambda)$ if and only if

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n}U^k(m,\lambda)}{n}\sum_{i=0}^{m}\Phi(m,i,\lambda)h_i=0, \tag{2.3}$$

which proves the lemma. □

Consider the following consequences of the assumptions above for further reasoning. Suppose that $\lambda\in\rho_{NS}(E-U(m,\lambda))$ and λ is right stable point of the monodromy operator, such that λ define eigenspace $N(E-U(m,\lambda))$, which coincides with the values set of operator $U_0(\lambda)x=\lim_{n\to\infty}U_n(m,\lambda)x$. This operator satisfies the following conditions [6]:

$$\text{(i)}\quad U_0(\lambda)=U_0^2(\lambda),\qquad \text{(ii)}\quad U_0(\lambda)=U(m,\lambda)U_0(\lambda),\qquad \text{(iii)}\quad U_0(\lambda)=U_0(\lambda)U(m,\lambda). \tag{2.4}$$

Lemma 2.2. *Operator $E-U(m,\lambda)+U_0(\lambda):\mathbf{B}\to\mathbf{B}$ has bounded inverse of the form*

$$(E-U(m,\lambda)+U_0(\lambda))^{-1}=\sum_{k=0}^{\infty}(\mu-1)^k\left\{\sum_{l=0}^{\infty}\mu^{-l-1}(U(m,\lambda)-U_0(\lambda))^l\right\}^{k+1}, \tag{2.5}$$

for all $\mu>1:|1-\mu|<1/\|R_\mu\|$.

Proof. Let us show that $\operatorname{Ker}(I-U(m,\lambda)+U_0(\lambda))=\overline{0}$. Indeed, if $x\in\operatorname{Ker}(I-U(m,\lambda)+U_0(\lambda))$, then

$$(I-U(m,\lambda)+U_0(\lambda))x=\overline{0}. \tag{2.6}$$

Since $(I-U(m,\lambda))x\in\operatorname{Im}(I-U(m,\lambda))$ and $U_0(\lambda)x\in\operatorname{Ker}(I-U(m,\lambda))$ [6], subspaces $\operatorname{Im}(I-U(m,\lambda))$ and $\operatorname{Ker}(I-U(m,\lambda))$ intersect only at zero point, and condition (2.6) is satisfied if and only if $(I-U(m,\lambda))x=\overline{0}$ and $U_0(\lambda)x=\overline{0}$. This is possible if and only if $x=\overline{0}$. Let us show that $\operatorname{Im}(I-U(m,\lambda)+U_0(\lambda))=\mathbf{B}$. Note [6] $\mathbf{B}=\operatorname{Ker}(I-U(m,\lambda))\oplus\operatorname{Im}(I-U(m,\lambda))=\operatorname{Im}(U_0(\lambda))\oplus Im(I-U(m,\lambda))$. It follows from the last decomposition that any element $x\in\mathbf{B}$ has the form

to Contents, page 1

$(I - U(m,\lambda))y + U_0(\lambda)z$, where $y, z \in \mathbf{B}$, which proves that $\operatorname{Im}(I - U(m,\lambda) + U_0(\lambda)) = \mathbf{B}$. Hence according to the Banach theorem [6] original operator has inverse since it bijectively maps $\mathbf{B}$ to itself. Therefore point $\mu = 1$ is regular [6] for the operator $\mu I - U(m,\lambda) + U_0(\lambda)$. Since powers of the operator $U(m,\lambda)$ are uniformly bounded and spectral radius $r_{U(m,\lambda)} \leqslant 1$ ($\sqrt[n]{\|U(m,\lambda)^n\|} \leqslant \sqrt[n]{c}$, then $r_{U(m,\lambda)} = \lim_{n\to\infty} \sqrt[n]{\|U(m,\lambda)^n\|} \leqslant \lim_{n\to\infty} \sqrt[n]{c} = 1$). It is well known [6] that resolvent set of a bounded operator is open. Number $\mu = 1 \in \rho(U(m,\lambda) - U_0(\lambda))$; thus there exist a neighborhood of μ such that each point from the neighborhood belongs to resolvent set. For any point $\mu > r_{(U(m,\lambda) - U_0(\lambda))}$ that belongs to the neighborhood there exists a resolvent [6], which has the form of converging in the norm series

$$R_\mu := R_\mu(U(m,\lambda) - U_0(\lambda)) = \sum_{l=0}^{\infty} \mu^{-l-1}(U(m,\lambda) - U_0(\lambda))^l. \tag{2.7}$$

Using the analyticity of the resolvent and well-known identity for points $\mu > 1$ such that $|1-\mu| < 1/(\|R_\mu(U(m,\lambda) - U_0(\lambda))\|)$, we obtain

$$R_1 = \sum_{k=0}^{\infty} (\mu - 1)^k R_\mu^{k+1}. \tag{2.8}$$

Finally, by substituting the series in the equation above, we get (2.5), which proves the lemma. □

Let us introduce some notation first before proving next statement.

Definition 2.3. Operator $L^- \in \mathcal{L}(\mathbf{B})$ is called generalized inverse for operator $L \in \mathcal{L}(\mathbf{B})$ [3] if the following conditions hold:

$$(1) \quad LL^-L = L, \qquad (2) \quad L^-LL^- = L^-. \tag{2.9}$$

Lemma 2.4. *Operator $E - U(m,\lambda)$ is generalized inverse and*

$$(E - U(m,\lambda))^- = (E - U(m,\lambda) + U_0(\lambda))^{-1} - U_0(\lambda), \tag{2.10}$$

or in the form of converging operator series

$$(E - U(m,\lambda))^- = \sum_{k=0}^{\infty} (\mu - 1)^k \left\{ \sum_{l=0}^{\infty} \mu^{-l-1} (U(m,\lambda) - U_0(\lambda))^l \right\}^{k+1} - U_0(\lambda), \tag{2.11}$$

for all $\mu > 1 : |1-\mu| < 1/\|R_\mu\|$.

to Contents, page 1

Proof. It suffices to check conditions (1) and (2) of the Definition 2.3. We use both representations (2.10), (2.11) and the expression (2.4) for operator $U_0(\lambda)$. Consider the following product:

$$\begin{aligned}
&(I-U(m,\lambda))\Big((I-U(m,\lambda)+U_0(\lambda))^{-1}-U_0(\lambda)\Big)(I-U(\lambda))\\
&\quad=((I-U(m,\lambda)+U_0(\lambda))\ -U_0(\lambda))\times\Big((I-U(m,\lambda)+U_0(\lambda))^{-1}-U_0(\lambda)\Big)(I-U(m,\lambda))\\
&\quad=\Big(I-U_0(\lambda)(I-U(m,\lambda)+U_0(\lambda))^{-1}-(I-U(m,\lambda)+U_0(\lambda))U_0(\lambda)+U_0(\lambda)^2\Big)\\
&\qquad\times(I-U(m,\lambda))\\
&\quad=\Big(I-U_0(\lambda)(I-U(m,\lambda)+U_0(\lambda))^{-1}\Big)\times(I-U(m,\lambda))\\
&\quad=\Big(I-U_0(\lambda)(I-U(m,\lambda)+U_0(\lambda))^{-1}\Big)((I-U(m,\lambda)+U_0(\lambda))-U_0(\lambda))\\
&\quad=I-U(m,\lambda)+U_0(\lambda)-U_0(\lambda)-U_0(\lambda)+U_0(\lambda)(I-U(m,\lambda)+U_0(\lambda))^{-1}U_0(\lambda)\\
&\quad=I-U(m,\lambda)-U_0(\lambda)+U_0(\lambda)(I-U(m,\lambda)+U_0(\lambda))^{-1}U_0(\lambda).
\end{aligned}\tag{2.12}$$

Note that $U_0(\lambda)(U(m,\lambda)-U_0(\lambda))^l=0$ for any $l\in\mathbb{N}$ (this directly follows from (2.4) using formula of binominal coefficient). Now, prove that

$$\begin{aligned}
U_0(\lambda)(I-U(m,\lambda)+U_0(\lambda))^{-1}U_0(\lambda)&=U_0(\lambda)(I-U(m,\lambda)+U_0(\lambda))^{-1}\\
&=(I-U(m,\lambda)+U_0(\lambda))^{-1}U_0(\lambda)\\
&=U_0(\lambda).
\end{aligned}\tag{2.13}$$

Indeed

$$\begin{aligned}
U_0(\lambda)(I-U(m,\lambda)+U_0(\lambda))^{-1}U_0(\lambda)&=\sum_{k=0}^{\infty}(\mu-1)^kU_0(\lambda)\left\{\sum_{l=0}^{\infty}\mu^{-l-1}(U(m,\lambda)-U_0(\lambda))^l\right\}^{k+1}U_0(\lambda)\\
&=\sum_{k=0}^{\infty}\Bigg(\left(\mu^{-1}\right)^{k+1}(\mu-1)^kU_0(\lambda)+(\mu-1)^kU_0(\lambda)\\
&\qquad\times\left\{\sum_{l=1}^{\infty}\mu^{-l-1}(U(m,\lambda)-U_0(\lambda))^l\right\}^{k+1}\Bigg)U_0(\lambda)\\
&=\sum_{k=0}^{+\infty}\mu^{-k-1}(\mu-1)^kU_0(\lambda)\\
&=\frac{1}{\mu}\sum_{k=0}^{+\infty}\left(\frac{\mu-1}{\mu}\right)^kU_0(\lambda)\\
&=\frac{1}{\mu}\frac{1}{1-(\mu-1)/\mu}U_0(\lambda)\\
&=U_0(\lambda).
\end{aligned}\tag{2.14}$$

to Contents, page 1

Thus

$$I - U(m,\lambda) - U_0(\lambda) + U_0(\lambda)(I - U(m,\lambda) + U_0(\lambda))^{-1}U_0(\lambda) = I - U(m,\lambda). \tag{2.15}$$

We have that the operator $I - U(m,\lambda)$ satisfies condition (1) of the Definition 2.3. Let us check condition (2)

$$\begin{aligned}
&\left((I - U(m,\lambda) + U_0(\lambda))^{-1} - U_0(\lambda)\right)(I - U(m,\lambda))\left((I - U(m,\lambda) + U_0(\lambda))^{-1} - U_0(\lambda)\right) \\
&\quad= \left((I - U(m,\lambda) + U_0(\lambda))^{-1} - U_0(\lambda)\right)((I - U(m,\lambda) + U_0(\lambda)) - U_0(\lambda)) \\
&\qquad\times \left((I - U(m,\lambda) + U_0(\lambda))^{-1} - U_0(\lambda)\right) \\
&\quad= \left(I - U_0(\lambda)(I - U(m,\lambda) + U_0(\lambda)) - (I - U(m,\lambda) + U_0(\lambda))^{-1}U_0(\lambda) + U_0(\lambda)^2\right) \\
&\qquad\times \left((I - U(m,\lambda) + U_0(\lambda))^{-1} - U_0(\lambda)\right) \\
&\quad= \left(I - (I - U(m,\lambda) + U_0(\lambda))^{-1}U_0(\lambda)\right)\left((I - U(m,\lambda) + U_0(\lambda))^{-1} - U_0(\lambda)\right) \\
&\quad= (I - U(m,\lambda) + U_0(\lambda))^{-1} - (I - U(m,\lambda) + U_0(\lambda))^{-1}U_0(\lambda)(I - U(m,\lambda) + U_0(\lambda))^{-1} \\
&\qquad - U_0(\lambda) + (I - U(m,\lambda) + U_0(\lambda))^{-1}U_0(\lambda) \\
&\quad= (I - U(m,\lambda) + U_0(\lambda))^{-1} - U_0(\lambda) - U_0(\lambda) + U_0(\lambda) \\
&\quad= (I - U(m,\lambda) + U_0(\lambda))^{-1} - U_0(\lambda).
\end{aligned} \tag{2.16}$$

□

3. Proof of Theorem 1.1

According to general theory of linear equations solvability [3], we obtain that the problem (1.1), (1.2) is solvable for sets $\{h_n\}_n \in \mathbb{Z}_+$ that satisfy the condition

$$U_0(\lambda)g(m,\lambda) = 0. \tag{3.1}$$

This condition along with Lemma 2.1 is equivalent to represantion (a) of the Theorem 1.1.

Under such a condition, all solutions of the problem (1.1), (1.2) have the form

$$\begin{aligned}
x_n &= U(n,\lambda)U_0(\lambda)c + U(n,\lambda)(I - U(m,\lambda))^{-}g(m,\lambda) + g(n,\lambda) \\
&= U(n,\lambda)U_0(\lambda)c + U(n,\lambda)\sum_{k=0}^{\infty}(\mu - 1)^k\left\{\sum_{l=0}^{\infty}\mu^{-l-1}(U(m,\lambda) - U_0(\lambda))^l\right\}^{k+1} g(m,\lambda) \\
&\quad - U(n,\lambda)U_0(\lambda)g(m,\lambda) + g(n,\lambda),
\end{aligned} \tag{3.2}$$

which along with notations introduced is equivalent to representation (b) of the theorem.

to Contents, page 1

4. Comments and Examples

Remark 4.1. Suppose B is Hilbert space, in such case we can show that formulas (2.10), (2.11) give us the representation for the Moore-Penrose pseudoinverse [7, 8] for $E - U(m,\lambda)$ with $U_0(\lambda)$ being self-adjoint operator (orthogonal projector) [6].

Remark 4.2. Supposing $A_k^{-1} \in \mathcal{L}(\mathbf{B}) \in L(B)$ exist for all $k = \overline{0, m-1}$, then the following equation holds: $\Phi(k,i,\lambda) = U(k,\lambda)U^{-1}(i,\lambda)$, $k > i$. This allows representing the solutions of (1.1), (1.2) using only the family of operators $U(n,\lambda)$ and their inverse.

Let us illustrate the statements proved above on example of two-dimensional systems.

(1) Consider equation

$$\overrightarrow{x}_{n+1} = \lambda A_{n+1}\overrightarrow{x}_n + \overrightarrow{h}_{n+1}, \quad n \geqslant 0 \tag{4.1}$$

with periodicity condition

$$\overrightarrow{x}_3 = \overrightarrow{x}_0, \tag{4.2}$$

where $\overrightarrow{x}_n = (x_n^1, x_n^2)^T$, $x_n^1, x_n^2 \in \mathbb{R}$, $\overrightarrow{h}_n = ((3\sqrt{3}r)/4\pi, 0)^T$,

$$A_n = \begin{pmatrix} -\dfrac{1}{2} & -\dfrac{\sqrt{3}}{2} \\ \dfrac{\sqrt{3}}{2} & -\dfrac{1}{2} \end{pmatrix}, \quad \forall n \geqslant 0. \tag{4.3}$$

It is easy to see that

$$\overrightarrow{x}_3 = \lambda^3\overrightarrow{x}_0 + g(3,\lambda), \tag{4.4}$$

where

$$g(3,\lambda) = \left(\frac{-3\sqrt{3}r\lambda - 3\sqrt{3}r\lambda^2 + 6\sqrt{3}r}{8\pi}, \frac{9r\lambda - 9r\lambda^2}{8\pi}\right)^T. \tag{4.5}$$

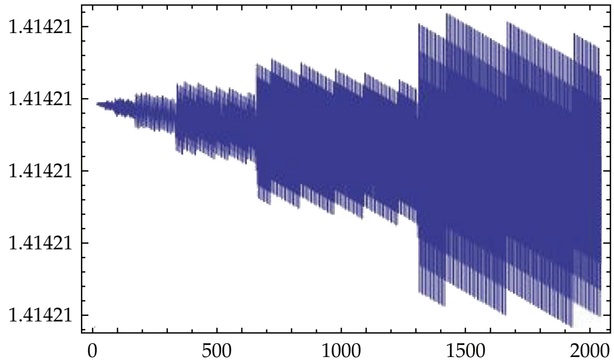
Then the following hold for all $k \geqslant 0$

$$U(3k+1,\lambda) = \lambda^{3k+1}A_2, \qquad U(3k+2,\lambda) = \lambda^{3k+2}\begin{pmatrix} -\dfrac{1}{2} & \dfrac{\sqrt{3}}{2} \\ \dfrac{\sqrt{3}}{2} & -\dfrac{1}{2} \end{pmatrix}, \qquad U(3k+3,\lambda) = \lambda^{3k+3}E. \tag{4.6}$$

By substituting periodicity condition (4.2) into (4.4) we obtain an equation depending on $\overrightarrow{x}_0$:

$$\left(1 - \lambda^3\right)\overrightarrow{x}_0 = g(3,\lambda). \tag{4.7}$$

to Contents, page 1



$r = -0.96, w = -0.25, c_1 = 1, c_2 = 1, T = 2000$

Figure 1

Consider the case when $\lambda = 1$. In such case (4.7) turns into $0\overrightarrow{x}_0 = (0,0)^T$ which holds for arbitrary initial vector $\overrightarrow{x}_0 \in \mathbb{R}^2$. Obviously $U^n(1,1) = U(n,1)$ and $U_0(1) = E$. According to Theorem 1.1, all periodic solutions of (4.1) have the form

$$\begin{pmatrix} x_n^1(c_1,c_2) \\ x_n^2(c_1,c_2) \end{pmatrix} = \begin{pmatrix} \cos\dfrac{2\pi}{3}n & \sin\dfrac{2\pi}{3}n \\ -\sin\dfrac{2\pi}{3}n & \cos\dfrac{2\pi}{3}n \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \dfrac{3r}{2\pi}\sin\dfrac{2\pi}{3}n \\ 0 \end{pmatrix}, \tag{4.8}$$

for all $\overrightarrow{c} = (c_1,c_2)^T \in \mathbb{R}^2$.

(2) We can search for periodic solutions of any period w in previous problem. They have common view

$$\overrightarrow{x}_n(c_1,c_2,w,r) = \begin{pmatrix} \cos\dfrac{2\pi}{w}n & \sin\dfrac{2\pi}{w}n \\ -\sin\dfrac{2\pi}{w}n & \cos\dfrac{2\pi}{w}n \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \dfrac{rw}{2\pi}\sin\dfrac{2\pi}{w}n \\ 0 \end{pmatrix}, \tag{4.9}$$

where c_1, c_2, w, r are parameters.

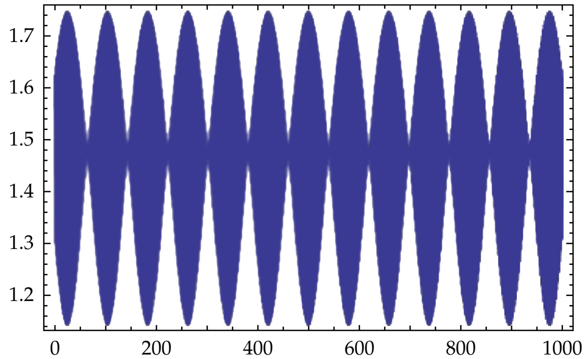
To illustrate complexity of the set we did the following.

Recall that the length of vector $\overrightarrow{x}_n$ is $\ell\overrightarrow{x}_n = \sqrt{(\overrightarrow{x}_n^1)^2 + (\overrightarrow{x}_n^2)^2}$. System (4.9) was implemented using the Wolfram Mathematica 7 framework. x-axis corresponds to time, while y-axis corresponds to the length of the vector. The length of the vector was calculated in the integer moments of time n. The points obtained in such way were connected in a piecewise linear way. The results obtained for particular values of the parameters are depicted on the following figures.

We can see how the trajectory of vector length densely fills rectangle or turns into a line (Figures 3 and 4). Figures 1, 2, 5, and 6 demonstrate that the trajectory can fill structured sets. The structure depicted on Figure 1 resembles fractal.

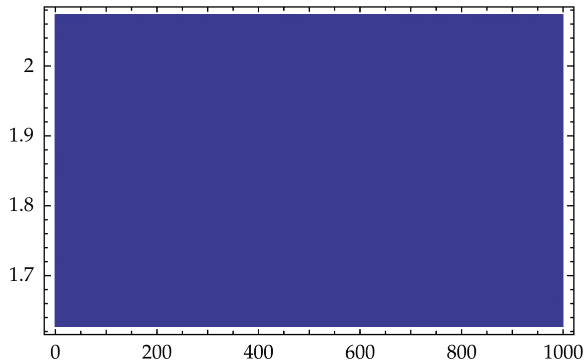
This allows us to conclude that behavior of the system is rather complex; it can undergo unpredictable changes with the slightest variations of a single parameter. We must admit that effects described need further theoretical investigation.

to Contents, page 1



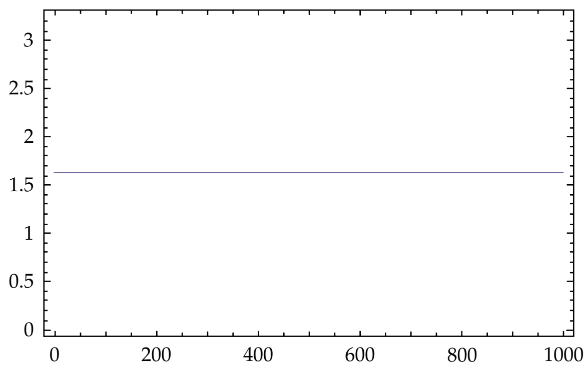
$r = -0.96, w = -3.95, c_1 = -1.23, c_2 = -1.06, T = 1000$

Figure 2



$r = -0.96, w = 4, c_1 = -1.23, c_2 = -1.06, T = 1000$

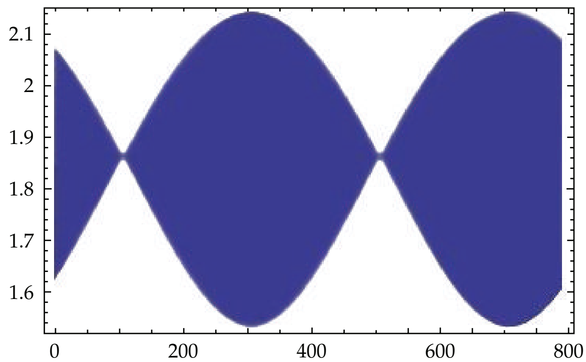
Figure 3



$r = -0.96, w = 0.5, c_1 = -1.23, c_2 = -1.06, T = 1000$

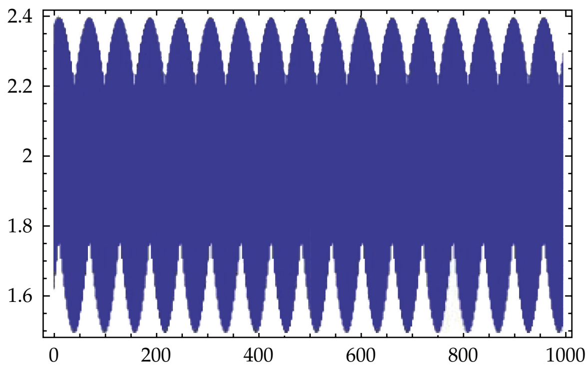
Figure 4

to Contents, page 1



$r = -0.96, w = 3.99, c_1 = -1.23, c_2 = -1.06, T = 790$

Figure 5



$r = -0.96, w = 5.9, c_1 = -1.23, c_2 = -1.06, T = 1000$

Figure 6

Acknowledgment

A. A. Boichuk was supported by the Grant 1/0090/09 of the Grant Agency of Slovak Republic (VEGA) and by the Project APVV-0700-07 of the Slovak Research and Development Agency.

References

[1] B. P. Demidovich, *Lectures Mathematical Theory of Stability*, Nauka, Moscow, Russia, 1967.
[2] Y. L. Daletskyi and M. G. Krein, *Stability of Solutions of Differential Equations in Banach Space*, Nauka, Moscow, Russia, 1970.
[3] A. A. Boichuk and A. M. Samoilenko, *Generalized Inverse Operators and Fredholm Boundary Value Problems*, VSP, Utrecht, The Netherlands, 2004.
[4] I. D. Chueshov, *Introduction to the Theory of Infinite—Dimensional Dissipative Systems*, ACTA, Kharkov, Ukraine, 2002.
[5] A. A. Boichuk and A. A. Pokutnyi, "Criterion of solvability of difference equations in Banach space," in *Bulgarian-Turkish-Ukrainian Scientific Conference Mathematical Analysis, Differential Equations and Their Applications*, pp. 241–247, Prof. Marin Drinov Academic Publishing House, Sunny Beach, Bulgaria, September 2010.

to Contents, page 1

[6] K. Iosida, *Functional Analysis*, Springer, Berlin, Germany, 1965.
[7] E. H. Moore, "On the reciprocal of the general algebraic matrix (Abstract)," *Bulletin of the American Mathematical Society*, no. 26, pp. 394–395, 1920.
[8] R. Penrose, "A generalized inverse for matrices," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 51, no. 3, pp. 406–413, 1955.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 390543, 16 pages
doi:10.1155/2011/390543

Research Article

Positive Solutions to Boundary Value Problems of Nonlinear Fractional Differential Equations

Yige Zhao,[1] Shurong Sun,[1,2] Zhenlai Han,[1] and Qiuping Li[1]

[1] *School of Science, University of Jinan, Jinan, Shandong 250022, China*
[2] *Department of Mathematics and Statistics, Missouri University of Science and Technology Rolla, MO 65409-0020, USA*

Correspondence should be addressed to Shurong Sun, sshrong@163.com

Received 23 September 2010; Revised 5 November 2010; Accepted 6 December 2010

Academic Editor: Josef Diblík

We study the existence of positive solutions for the boundary value problem of nonlinear fractional differential equations $D_{0^+}^{\alpha}u(t) + \lambda f(u(t)) = 0$, $0 < t < 1$, $u(0) = u(1) = u'(0) = 0$, where $2 < \alpha \leq 3$ is a real number, $D_{0^+}^{\alpha}$ is the Riemann-Liouville fractional derivative, λ is a positive parameter, and $f : (0, +\infty) \to (0, +\infty)$ is continuous. By the properties of the Green function and Guo-Krasnosel'skii fixed point theorem on cones, the eigenvalue intervals of the nonlinear fractional differential equation boundary value problem are considered, some sufficient conditions for the nonexistence and existence of at least one or two positive solutions for the boundary value problem are established. As an application, some examples are presented to illustrate the main results.

1. Introduction

Fractional differential equations have been of great interest recently. It is caused both by the intensive development of the theory of fractional calculus itself and by the applications; see [1–4]. It should be noted that most of papers and books on fractional calculus are devoted to the solvability of linear initial fractional differential equations on terms of special functions.

Recently, there are some papers dealing with the existence of solutions (or positive solutions) of nonlinear initial fractional differential equations by the use of techniques of nonlinear analysis (fixed-point theorems, Leray-Schauder theory, Adomian decomposition method, etc.); see [5–11]. In fact, there has the same requirements for boundary conditions. However, there exist some papers considered the boundary value problems of fractional differential equations; see [12–19].

to Contents, page 1

Yu and Jiang [19] examined the existence of positive solutions for the following problem:

$$\begin{gathered}D_{0^+}^{\alpha}u(t)+f(t,u(t))=0, \quad 0<t<1,\\ u(0)=u(1)=u'(0)=0,\end{gathered} \tag{1.1}$$

where $2<\alpha\leq 3$ is a real number, $f\in C([0,1]\times[0,+\infty),(0,+\infty))$, and $D_{0^+}^{\alpha}$ is the Riemann-Liouville fractional differentiation. By using the properties of the Green function, they obtained some existence criteria for one or two positive solutions for singular and nonsingular boundary value problems by means of the Krasnosel'skii fixed point theorem and a mixed monotone method.

To the best of our knowledge, there is very little known about the existence of positive solutions for the following problem:

$$\begin{gathered}D_{0^+}^{\alpha}u(t)+\lambda f(u(t))=0, \quad 0<t<1,\\ u(0)=u(1)=u'(0)=0,\end{gathered} \tag{1.2}$$

where $2<\alpha\leq 3$ is a real number, $D_{0^+}^{\alpha}$ is the Riemann-Liouville fractional derivative, λ is a positive parameter and $f:(0,+\infty)\rightarrow(0,+\infty)$ is continuous.

On one hand, the boundary value problem in [19] is the particular case of problem (1.2) as the case of $\lambda=1$. On the other hand, as Yu and Jiang discussed in [19], we also give some existence results by the fixed point theorem on a cone in this paper. Moreover, the purpose of this paper is to derive a λ-interval such that, for any λ lying in this interval, the problem (1.2) has existence and multiplicity on positive solutions.

In this paper, by analogy with boundary value problems for differential equations of integer order, we firstly give the corresponding Green function named by fractional Green's function and some properties of the Green function. Consequently, the problem (1.2) is reduced to an equivalent Fredholm integral equation. Finally, by the properties of the Green function and Guo-Krasnosel'skii fixed point theorem on cones, the eigenvalue intervals of the nonlinear fractional differential equation boundary value problem are considered, some sufficient conditions for the nonexistence and existence of at least one or two positive solutions for the boundary value problem are established. As an application, some examples are presented to illustrate the main results.

2. Preliminaries

For the convenience of the reader, we give some background materials from fractional calculus theory to facilitate analysis of problem (1.2). These materials can be found in the recent literature; see [19–21].

Definition 2.1 (see [20]). The Riemann-Liouville fractional derivative of order $\alpha>0$ of a continuous function $f:(0,+\infty)\rightarrow\mathbb{R}$ is given by

$$D_{0^+}^{\alpha}f(t)=\frac{1}{\Gamma(n-\alpha)}\left(\frac{d}{dt}\right)^{(n)}\int_0^t\frac{f(s)}{(t-s)^{\alpha-n+1}}ds, \tag{2.1}$$

to Contents, page 1

where $n = [\alpha] + 1$, $[\alpha]$ denotes the integer part of number α, provided that the right side is pointwise defined on $(0, +\infty)$.

Definition 2.2 (see [20]). The Riemann-Liouville fractional integral of order $\alpha > 0$ of a function $f : (0, +\infty) \to \mathbb{R}$ is given by

$$I_{0^+}^{\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s) ds, \tag{2.2}$$

provided that the right side is pointwise defined on $(0, +\infty)$.

From the definition of the Riemann-Liouville derivative, we can obtain the following statement.

Lemma 2.3 (see [20]). *Let $\alpha > 0$. If we assume $u \in C(0,1) \cap L(0,1)$, then the fractional differential equation*

$$D_{0^+}^{\alpha} u(t) = 0 \tag{2.3}$$

has $u(t) = c_1 t^{\alpha-1} + c_2 t^{\alpha-2} + \cdots + c_N t^{\alpha-N}$, $c_i \in \mathbb{R}$, $i = 1, 2, \ldots, N$, as unique solutions, where N is the smallest integer greater than or equal to α.

Lemma 2.4 (see [20]). *Assume that $u \in C(0,1) \cap L(0,1)$ with a fractional derivative of order $\alpha > 0$ that belongs to $C(0,1) \cap L(0,1)$. Then*

$$I_{0^+}^{\alpha} D_{0^+}^{\alpha} u(t) = u(t) + c_1 t^{\alpha-1} + c_2 t^{\alpha-2} + \cdots + c_N t^{\alpha-N}, \tag{2.4}$$

for some $c_i \in \mathbb{R}$, $i = 1, 2, \ldots, N$, where N is the smallest integer greater than or equal to α.

In the following, we present the Green function of fractional differential equation boundary value problem.

Lemma 2.5 (see [19]). *Let $h \in C[0,1]$ and $2 < \alpha \leq 3$. The unique solution of problem*

$$\begin{gathered} D_{0^+}^{\alpha} u(t) + h(t) = 0, \quad 0 < t < 1, \\ u(0) = u(1) = u'(0) = 0 \end{gathered} \tag{2.5}$$

is

$$u(t) = \int_0^1 G(t,s) h(s) ds, \tag{2.6}$$

to Contents, page 1

where

$$G(t,s)=\begin{cases}\dfrac{t^{\alpha-1}(1-s)^{\alpha-1}-(t-s)^{\alpha-1}}{\Gamma(\alpha)}, & 0\le s\le t\le 1,\\ \dfrac{t^{\alpha-1}(1-s)^{\alpha-1}}{\Gamma(\alpha)}, & 0\le t\le s\le 1.\end{cases} \tag{2.7}$$

Here $G(t,s)$ *is called the Green function of boundary value problem* (2.5).

The following properties of the Green function play important roles in this paper.

Lemma 2.6 (see [19]). *The function* $G(t,s)$ *defined by* (2.7) *satisfies the following conditions:*

(1) $G(t,s)=G(1-s,1-t)$, *for* $t,s\in(0,1)$;

(2) $t^{\alpha-1}(1-t)s(1-s)^{\alpha-1}\le\Gamma(\alpha)G(t,s)\le(\alpha-1)s(1-s)^{\alpha-1}$, *for* $t,s\in(0,1)$;

(3) $G(t,s)>0$, *for* $t,s\in(0,1)$;

(4) $t^{\alpha-1}(1-t)s(1-s)^{\alpha-1}\le\Gamma(\alpha)G(t,s)\le(\alpha-1)(1-t)t^{\alpha-1}$, *for* $t,s\in(0,1)$.

The following lemma is fundamental in the proofs of our main results.

Lemma 2.7 (see [21]). *Let* X *be a Banach space, and let* $P\subset X$ *be a cone in* X. *Assume* $\Omega_1,\ \Omega_2$ *are open subsets of* X *with* $0\in\Omega_1\subset\overline{\Omega}_1\subset\Omega_2$, *and let* $S:P\to P$ *be a completely continuous operator such that, either*

$(A1)$ $\|Sw\|\le\|w\|$, $w\in P\cap\partial\Omega_1$, $\|Sw\|\ge\|w\|$, $w\in P\cap\partial\Omega_2$ *or*

$(A2)$ $\|Sw\|\ge\|w\|$, $w\in P\cap\partial\Omega_1$, $\|Sw\|\le\|w\|$, $w\in P\cap\partial\Omega_2$.

Then S *has a fixed point in* $P\cap(\overline{\Omega}_2\setminus\Omega_1)$.

For convenience, we set $q(t)=t^{\alpha-1}(1-t), k(s)=s(1-s)^{\alpha-1}$; then

$$q(t)k(s)\le\Gamma(\alpha)G(t,s)\le(\alpha-1)k(s). \tag{2.8}$$

3. Main Results

In this section, we establish the existence of positive solutions for boundary value problem (1.2).

Let Banach space $E=C[0,1]$ be endowed with the norm $\|u\|=\max_{0\le t\le 1}|u(t)|$. Define the cone $P\subset E$ by

$$P=\left\{u\in E: u(t)\ge\frac{q(t)}{\alpha-1}\|u\|,\ t\in[0,1]\right\}. \tag{3.1}$$

Suppose that u is a solution of boundary value problem (1.2). Then

$$u(t)=\lambda\int_0^1 G(t,s)f(u(s))ds,\quad t\in[0,1]. \tag{3.2}$$

to Contents, page 1

We define an operator $A_\lambda : P \to E$ as follows:

$$(A_\lambda u)(t) = \lambda \int_0^1 G(t,s) f(u(s)) ds, \quad t \in [0,1]. \tag{3.3}$$

By Lemma 2.6, we have

$$\begin{aligned} \|A_\lambda u\| &\leq \frac{\lambda}{\Gamma(\alpha)} \int_0^1 (\alpha - 1) k(s) f(u(s)) ds, \\ (A_\lambda u)(t) &\geq \frac{\lambda}{\Gamma(\alpha)} \int_0^1 q(t) k(s) f(u(s)) ds \\ &\geq \frac{q(t)}{\alpha - 1} \|A_\lambda u\|. \end{aligned} \tag{3.4}$$

Thus, $A_\lambda(P) \subset P$.

Then we have the following lemma.

Lemma 3.1. *$A_\lambda : P \to P$ is completely continuous.*

Proof. The operator $A_\lambda : P \to P$ is continuous in view of continuity of $G(t,s)$ and $f(u(t))$. By means of the Arzela-Ascoli theorem, $A_\lambda : P \to P$ is completely continuous.

For convenience, we denote

$$\begin{gathered} F_0 = \lim_{u \to 0^+} \sup \frac{f(u)}{u}, \qquad F_\infty = \lim_{u \to +\infty} \sup \frac{f(u)}{u}, \\ f_0 = \lim_{u \to 0^+} \inf \frac{f(u)}{u}, \qquad f_\infty = \lim_{u \to +\infty} \inf \frac{f(u)}{u}, \\ C_1 = \frac{1}{\Gamma(\alpha)} \int_0^1 (\alpha - 1) k(s) ds, \\ C_2 = \frac{1}{\Gamma(\alpha)} \int_0^1 \frac{1}{(\alpha - 1)} q(s) k(s) ds, \\ C_3 = \frac{1}{\Gamma(\alpha)} \int_0^1 \frac{1}{(\alpha - 1)} k(s) ds. \end{gathered} \tag{3.5}$$

□

Theorem 3.2. *If there exists $l \in (0,1)$ such that $q(l) f_\infty C_2 > F_0 C_1$ holds, then for each*

$$\lambda \in \left((q(l) f_\infty C_2)^{-1}, (F_0 C_1)^{-1} \right), \tag{3.6}$$

the boundary value problem (1.2) *has at least one positive solution. Here we impose $(q(l) f_\infty C_2)^{-1} = 0$ if $f_\infty = +\infty$ and $(F_0 C_1)^{-1} = +\infty$ if $F_0 = 0$.*

to Contents, page 1

Proof. Let λ satisfy (3.6) and $\varepsilon > 0$ be such that

$$\left(q(l)\left(f_\infty - \varepsilon\right)C_2\right)^{-1} \le \lambda \le ((F_0 + \varepsilon)C_1)^{-1}. \tag{3.7}$$

By the definition of F_0, we see that there exists $r_1 > 0$ such that

$$f(u) \le (F_0 + \varepsilon)u, \quad \text{for } 0 < u \le r_1. \tag{3.8}$$

So if $u \in P$ with $\|u\| = r_1$, then by (3.7) and (3.8), we have

$$\begin{aligned} \|A_\lambda u\| &\le \frac{\lambda}{\Gamma(\alpha)} \int_0^1 (\alpha - 1)k(s)f(u(s))ds \\ &\le \frac{\lambda}{\Gamma(\alpha)} \int_0^1 (\alpha - 1)k(s)(F_0 + \varepsilon)r_1 ds \\ &= \lambda(F_0 + \varepsilon)r_1 C_1 \\ &\le r_1 = \|u\|. \end{aligned} \tag{3.9}$$

Hence, if we choose $\Omega_1 = \{u \in E : \|u\| < r_1\}$, then

$$\|A_\lambda u\| \le \|u\|, \quad \text{for } u \in P \cap \partial\Omega_1. \tag{3.10}$$

Let $r_3 > 0$ be such that

$$f(u) \ge \left(f_\infty - \varepsilon\right)u, \quad \text{for } u \ge r_3. \tag{3.11}$$

If $u \in P$ with $\|u\| = r_2 = \max\{2r_1, r_3\}$, then by (3.7) and (3.11), we have

$$\begin{aligned} \|A_\lambda u\| &\ge A_\lambda u(l) \\ &= \lambda \int_0^1 G(l,s)f(u(s))ds \\ &\ge \frac{\lambda}{\Gamma(\alpha)} \int_0^1 q(l)k(s)f(u(s))ds \\ &\ge \frac{\lambda}{\Gamma(\alpha)} \int_0^1 q(l)k(s)\left(f_\infty - \varepsilon\right)u(s)ds \\ &\ge \frac{\lambda}{\Gamma(\alpha)} \int_0^1 \frac{q(l)}{\alpha - 1} q(s)k(s)\left(f_\infty - \varepsilon\right)\|u\|ds \\ &= \lambda q(l)C_2\left(f_\infty - \varepsilon\right)\|u\| \ge \|u\|. \end{aligned} \tag{3.12}$$

to Contents, page 1

Thus, if we set $\Omega_2 = \{u \in E : \|u\| < r_2\}$, then

$$\|A_\lambda u\| \geq \|u\|, \quad \text{for } u \in P \cap \partial\Omega_2. \tag{3.13}$$

Now, from (3.10), (3.13), and Lemma 2.7, we guarantee that A_λ has a fixed-point $u \in P \cap (\overline{\Omega}_2 \setminus \Omega_1)$ with $r_1 \leq \|u\| \leq r_2$, and clearly u is a positive solution of (1.2). The proof is complete. □

Theorem 3.3. *If there exists $l \in (0,1)$ such that $q(l)C_2 f_0 > F_\infty C_1$ holds, then for each*

$$\lambda \in \left((q(l) f_0 C_2)^{-1}, (F_\infty C_1)^{-1} \right), \tag{3.14}$$

the boundary value problem (1.2) *has at least one positive solution. Here we impose $(q(l) f_0 C_2)^{-1} = 0$ if $f_0 = +\infty$ and $(F_\infty C_1)^{-1} = +\infty$ if $F_\infty = 0$.*

Proof. Let λ satisfy (3.14) and $\varepsilon > 0$ be such that

$$(q(l)(f_0 - \varepsilon)C_2)^{-1} \leq \lambda \leq ((F_\infty + \varepsilon)C_1)^{-1}. \tag{3.15}$$

From the definition of f_0, we see that there exists $r_1 > 0$ such that

$$f(u) \geq (f_0 - \varepsilon)u, \quad \text{for } 0 < u \leq r_1. \tag{3.16}$$

Further, if $u \in P$ with $\|u\| = r_1$, then similar to the second part of Theorem 3.2, we can obtain that $\|A_\lambda u\| \geq \|u\|$. Thus, if we choose $\Omega_1 = \{u \in E : \|u\| < r_1\}$, then

$$\|A_\lambda u\| \geq \|u\|, \quad \text{for } u \in P \cap \partial\Omega_2. \tag{3.17}$$

Next, we may choose $R_1 > 0$ such that

$$f(u) \leq (F_\infty + \varepsilon)u, \quad \text{for } u \geq R_1. \tag{3.18}$$

We consider two cases.

Case 1. Suppose f is bounded. Then there exists some $M > 0$, such that

$$f(u) \leq M, \quad \text{for } u \in (0, +\infty). \tag{3.19}$$

to Contents, page 1

We define $r_3 = \max\{2r_1, \lambda M C_1\}$, and $u \in P$ with $\|u\| = r_3$, then

$$\begin{aligned} \|A_\lambda u\| &\leq \frac{\lambda}{\Gamma(\alpha)} \int_0^1 (\alpha-1)k(s)f(u(s))ds \\ &\leq \frac{\lambda M}{\Gamma(\alpha)} \int_0^1 (\alpha-1)k(s)ds \\ &\leq \lambda M C_1 \\ &\leq r_3 \leq \|u\|. \end{aligned} \tag{3.20}$$

Hence,

$$\|A_\lambda u\| \leq \|u\|, \quad \text{for } u \in P_{r_3} = \{u \in P : \|u\| \leq r_3\}. \tag{3.21}$$

Case 2. Suppose f is unbounded. Then there exists some $r_4 > \max\{2r_1, R_1\}$, such that

$$f(u) \leq f(r_4), \quad \text{for } 0 < u \leq r_4. \tag{3.22}$$

Let $u \in P$ with $\|u\| = r_4$. Then by (3.15) and (3.18), we have

$$\begin{aligned} \|A_\lambda u\| &\leq \frac{\lambda}{\Gamma(\alpha)} \int_0^1 (\alpha-1)k(s)f(u(s))ds \\ &\leq \frac{\lambda}{\Gamma(\alpha)} \int_0^1 (\alpha-1)k(s)(F_\infty+\varepsilon)\|u\|ds \\ &\leq \lambda C_1(F_\infty+\varepsilon)\|u\| \\ &\leq \|u\|. \end{aligned} \tag{3.23}$$

Thus, (3.21) is also true.

In both Cases 1 and 2, if we set $\Omega_2 = \{u \in E : \|u\| < r_2 = \max\{r_3, r_4\}\}$, then

$$\|A_\lambda u\| \leq \|u\|, \quad \text{for } u \in P \cap \partial\Omega_2. \tag{3.24}$$

Now that we obtain (3.17) and (3.24), it follows from Lemma 2.7 that A_λ has a fixed-point $u \in P \cap (\overline{\Omega}_2 \setminus \Omega_1)$ with $r_1 \leq \|u\| \leq r_2$. It is clear u is a positive solution of (1.2). The proof is complete. □

Theorem 3.4. *Suppose there exist $l \in (0,1)$, $r_2 > r_1 > 0$ such that $q(l) > (\alpha-1)r_1/r_2$, and f satisfy*

$$\min_{(q(l)/(\alpha-1))r_1 \leq u \leq r_1} f(u) \geq \frac{r_1}{\lambda(\alpha-1)q(l)C_3}, \qquad \max_{0 \leq u \leq r_2} f(u) \leq \frac{r_2}{\lambda C_1}. \tag{3.25}$$

Then the boundary value problem (1.2) *has a positive solution $u \in P$ with $r_1 \leq \|u\| \leq r_2$.*

to Contents, page 1

Proof. Choose $\Omega_1 = \{u \in E : \|u\| < r_1\}$; then for $u \in P \cap \partial\Omega_1$, we have

$$\begin{aligned}
\|A_\lambda u\| &\geq A_\lambda u(l) \\
&= \lambda \int_0^1 G(l,s) f(u(s)) ds \\
&\geq \frac{\lambda}{\Gamma(\alpha)} \int_0^1 q(l)k(s) f(u(s)) ds \\
&\geq \frac{\lambda}{\Gamma(\alpha)} \int_0^1 q(l)k(s) \min_{(q(l)/(\alpha-1))r_1 \leq u \leq r_1} f(u(s)) ds \\
&\geq \lambda(\alpha-1)q(l)C_3 \frac{r_1}{\lambda(\alpha-1)q(l)C_3} \\
&= r_1 = \|u\|.
\end{aligned} \tag{3.26}$$

On the other hand, choose $\Omega_2 = \{u \in E : \|u\| < r_2\}$, then for $u \in P \cap \partial\Omega_2$, we have

$$\begin{aligned}
\|A_\lambda u\| &\leq \frac{\lambda}{\Gamma(\alpha)} \int_0^1 (\alpha-1)k(s) f(u(s)) ds \\
&\leq \frac{\lambda}{\Gamma(\alpha)} \int_0^1 (\alpha-1)k(s) \max_{0 \leq u \leq r_2} f(u(s)) ds \\
&\leq \lambda C_1 \frac{r_2}{\lambda C_1} \\
&= r_2 = \|u\|.
\end{aligned} \tag{3.27}$$

Thus, by Lemma 2.7, the boundary value problem (1.2) has a positive solution $u \in P$ with $r_1 \leq \|u\| \leq r_2$. The proof is complete. □

For the reminder of the paper, we will need the following condition.

(H) $(\min_{u \in [(q(l)/(\alpha-1))r,\, r]} f(u))/r > 0$, where $l \in (0,1)$.

Denote

$$\lambda_1 = \sup_{r>0} \frac{r}{C_1 \max_{0 \leq u \leq r} f(u)}, \tag{3.28}$$

$$\lambda_2 = \inf_{r>0} \frac{r}{C_3 \min_{(q(l)/(\alpha-1))r \leq u \leq r} f(u)}. \tag{3.29}$$

In view of the continuity of $f(u)$ and (H), we have $0 < \lambda_1 \leq +\infty$ and $0 \leq \lambda_2 < +\infty$.

Theorem 3.5. *Assume* (H) *holds. If* $f_0 = +\infty$ *and* $f_\infty = +\infty$, *then the boundary value problem* (1.2) *has at least two positive solutions for each* $\lambda \in (0, \lambda_1)$.

to Contents, page 1

Proof. Define

$$a(r) = \frac{r}{C_1 \max_{0\le u\le r} f(u)}. \tag{3.30}$$

By the continuity of $f(u)$, $f_0 = +\infty$ and $f_\infty = +\infty$, we have that $a(r) : (0,+\infty) \to (0,+\infty)$ is continuous and

$$\lim_{r\to 0} a(r) = \lim_{r\to+\infty} a(r) = 0. \tag{3.31}$$

By (3.28), there exists $r_0 \in (0,+\infty)$, such that

$$a(r_0) = \sup_{r>0} a(r) = \lambda_1; \tag{3.32}$$

then for $\lambda \in (0,\lambda_1)$, there exist constants c_1, c_2 $(0 < c_1 < r_0 < c_2 < +\infty)$ with

$$a(c_1) = a(c_2) = \lambda. \tag{3.33}$$

Thus,

$$f(u) \le \frac{c_1}{\lambda C_1}, \quad \text{for } u \in [0, c_1], \tag{3.34}$$

$$f(u) \le \frac{c_2}{\lambda C_1}, \quad \text{for } u \in [0, c_2]. \tag{3.35}$$

On the other hand, applying the conditions $f_0 = +\infty$ and $f_\infty = +\infty$, there exist constants d_1, d_2 $(0 < d_1 < c_1 < r_0 < c_2 < d_2 < +\infty)$ with

$$\frac{f(u)}{u} \ge \frac{1}{q^2(l)\lambda C_3}, \quad \text{for } u \in (0, d_1) \cup \left(\frac{q(l)}{\alpha-1}d_2, +\infty\right). \tag{3.36}$$

Then

$$\min_{(q(l)/(\alpha-1))d_1 \le u \le d_1} f(u) \ge \frac{d_1}{\lambda(\alpha-1)q(l)C_3}, \tag{3.37}$$

$$\min_{(q(l)/(\alpha-1))d_2 \le u \le d_2} f(u) \ge \frac{d_2}{\lambda(\alpha-1)q(l)C_3}. \tag{3.38}$$

By (3.34) and (3.37), (3.35) and (3.38), combining with Theorem 3.4 and Lemma 2.7, we can complete the proof. □

Corollary 3.6. *Assume* (H) *holds. If* $f_0 = +\infty$ *or* $f_\infty = +\infty$, *then the boundary value problem* (1.2) *has at least one positive solution for each* $\lambda \in (0, \lambda_1)$.

to Contents, page 1

Theorem 3.7. *Assume* (H) *holds. If* $f_0 = 0$ *and* $f_\infty = 0$*, then for each* $\lambda \in (\lambda_2, +\infty)$*, the boundary value problem* (1.2) *has at least two positive solutions.*

Proof. Define

$$b(r) = \frac{r}{C_3 \min_{(q(l)/(\alpha-1))r \le u \le r} f(u)}. \tag{3.39}$$

By the continuity of $f(u)$, $f_0 = 0$ and $f_\infty = 0$, we easily see that $b(r) : (0, +\infty) \to (0, +\infty)$ is continuous and

$$\lim_{r \to 0} b(r) = \lim_{r \to +\infty} b(r) = +\infty. \tag{3.40}$$

By (3.29), there exists $r_0 \in (0, +\infty)$, such that

$$b(r_0) = \inf_{r>0} b(r) = \lambda_2. \tag{3.41}$$

For $\lambda \in (\lambda_2, +\infty)$, there exist constants d_1, d_2 $(0 < d_1 < r_0 < d_2 < +\infty)$ with

$$b(d_1) = b(d_2) = \lambda. \tag{3.42}$$

Therefore,

$$\begin{aligned} f(u) &\ge \frac{d_1}{\lambda(\alpha-1)q(l)C_3}, \quad \text{for } u \in \left[\frac{q(l)}{\alpha-1}d_1, d_1\right], \\ f(u) &\ge \frac{d_2}{\lambda(\alpha-1)q(l)C_3}, \quad \text{for } u \in \left[\frac{q(l)}{\alpha-1}d_2, d_2\right]. \end{aligned} \tag{3.43}$$

On the other hand, using $f_0 = 0$, we know that there exists a constant c_1 $(0 < c_1 < d_1)$ with

$$\frac{f(u)}{u} \le \frac{1}{\lambda C_1}, \quad \text{for } u \in (0, c_1), \tag{3.44}$$

$$\max_{0 \le u \le c_1} f(u) \le \frac{c_1}{\lambda C_1}. \tag{3.45}$$

In view of $f_\infty = 0$, there exists a constant $c_2 \in (d_2, +\infty)$ such that

$$\frac{f(u)}{u} \le \frac{1}{\lambda C_1}, \quad \text{for } u \in (c_2, +\infty). \tag{3.46}$$

Let

$$M = \max_{0 \le u \le c_2} f(u), \quad c_2 \ge \lambda C_1 M. \tag{3.47}$$

to Contents, page 1

It is easily seen that

$$\max_{0\le u\le c_2} f(u) \le \frac{c_2}{\lambda C_1}. \tag{3.48}$$

By (3.45) and (3.48), combining with Theorem 3.4 and Lemma 2.7, the proof is complete. □

Corollary 3.8. *Assume* (H) *holds. If* $f_0 = 0$ *or* $f_\infty = 0$*, then for each* $\lambda \in (\lambda_2, +\infty)$*, the boundary value problem* (1.2) *has at least one positive solution.*

By the above theorems, we can obtain the following results.

Corollary 3.9. *Assume* (H) *holds. If* $f_0 = +\infty$*,* $f_\infty = d$*, or* $f_\infty = +\infty$*,* $f_0 = d$*, then for any* $\lambda \in (0, (dC_1)^{-1})$*, the boundary value problem* (1.2) *has at least one positive solution.*

Corollary 3.10. *Assume* (H) *holds. If* $f_0 = 0, f_\infty = d$*, or if* $f_\infty = 0$*,* $f_0 = d$*, then for any* $\lambda \in ((q(l)dC_2)^{-1}, +\infty)$*, the boundary value problem* (1.2) *has at least one positive solution.*

Remark 3.11. For the integer derivative case $\alpha = 3$, Theorems 3.2–3.7 also hold; we can find the corresponding existence results in [22].

4. Nonexistence

In this section, we give some sufficient conditions for the nonexistence of positive solution to the problem (1.2).

Theorem 4.1. *Assume* (H) *holds. If* $F_0 < +\infty$ *and* $F_\infty < \infty$*, then there exists a* $\lambda_0 > 0$ *such that for all* $0 < \lambda < \lambda_0$*, the boundary value problem* (1.2) *has no positive solution.*

Proof. Since $F_0 < +\infty$ and $F_\infty < +\infty$, there exist positive numbers m_1, m_2, r_1, and r_2, such that $r_1 < r_2$ and

$$\begin{aligned} f(u) &\le m_1 u, \quad \text{for } u \in [0, r_1], \\ f(u) &\le m_2 u, \quad \text{for } u \in [r_2, +\infty). \end{aligned} \tag{4.1}$$

Let $m = \max\{m_1, m_2, \max_{r_1\le u\le r_2}\{f(u)/u\}\}$. Then we have

$$f(u) \le mu, \quad \text{for } u \in [0, +\infty). \tag{4.2}$$

Assume $v(t)$ is a positive solution of (1.2). We will show that this leads to a contradiction for $0 < \lambda < \lambda_0 := (mC_1)^{-1}$. Since $A_\lambda v(t) = v(t)$ for $t \in [0,1]$,

$$\|v\| = \|A_\lambda v\| \le \frac{\lambda}{\Gamma(\alpha)} \int_0^1 (\alpha-1)k(s) f(v(s))\,ds \le \frac{m\lambda}{\Gamma(\alpha)} \|v\| \int_0^1 (\alpha-1)k(s)\,ds < \|v\|, \tag{4.3}$$

which is a contradiction. Therefore, (1.2) has no positive solution. The proof is complete. □

to Contents, page 1

Theorem 4.2. *Assume* (H) *holds. If* $f_0 > 0$ *and* $f_\infty > 0$*, then there exists a* $\lambda_0 > 0$ *such that for all* $\lambda > \lambda_0$*, the boundary value problem* (1.2) *has no positive solution.*

Proof. By $f_0 > 0$ and $f_\infty > 0$, we know that there exist positive numbers $n_1,\ n_2,\ r_1$, and r_2, such that $r_1 < r_2$ and

$$\begin{aligned} f(u) &\geq n_1 u, \quad \text{for } u \in [0, r_1], \\ f(u) &\geq n_2 u, \quad \text{for } u \in [r_2, +\infty). \end{aligned} \tag{4.4}$$

Let $n = \min\{n_1, n_2, \min_{r_1 \leq u \leq r_2}\{f(u)/u\}\} > 0$. Then we get

$$f(u) \geq nu, \quad \text{for } u \in [0, +\infty). \tag{4.5}$$

Assume $v(t)$ is a positive solution of (1.2). We will show that this leads to a contradiction for $\lambda > \lambda_0 := (q(l)nC_2)^{-1}$. Since $A_\lambda v(t) = v(t)$ for $t \in [0,1]$,

$$\|v\| = \|A_\lambda v\| \geq \frac{\lambda}{\Gamma(\alpha)} \int_0^1 q(l)k(s)f(v(s))ds > \|v\|, \tag{4.6}$$

which is a contradiction. Thus, (1.2) has no positive solution. The proof is complete. □

5. Examples

In this section, we will present some examples to illustrate the main results.

Example 5.1. Consider the boundary value problem

$$\begin{gathered} D_{0^+}^{5/2} u(t) + \lambda u^a = 0, \quad 0 < t < 1, \ a > 1, \\ u(0) = u(1) = u'(0) = 0. \end{gathered} \tag{5.1}$$

Since $\alpha = 5/2$, we have

$$\begin{aligned} C_1 &= \frac{1}{\Gamma(\alpha)} \int_0^1 (\alpha - 1)k(s)ds = \frac{1}{\Gamma(5/2)} \int_0^1 \frac{3}{2} s(1-s)^{3/2} ds = 0.1290, \\ C_2 &= \frac{1}{\Gamma(\alpha)} \int_0^1 \frac{1}{(\alpha-1)} q(s)k(s)ds = \frac{1}{\Gamma(5/2)} \int_0^1 \frac{2}{3} s^{5/2}(1-s)^{5/2} ds = 0.0077. \end{aligned} \tag{5.2}$$

Let $f(u) = u^a,\ a > 1$. Then we have $F_0 = 0$, $f_\infty = +\infty$. Choose $l = 1/2$. Then $q(1/2) = \sqrt{2}/8 = 0.1768$. So $q(l)C_2 f_\infty > F_0 C_1$ holds. Thus, by Theorem 3.2, the boundary value problem (5.1) has a positive solution for each $\lambda \in (0, +\infty)$.

to Contents, page 1

Example 5.2. Discuss the boundary value problem

$$D_{0^+}^{5/2}u(t) + \lambda u^b = 0, \quad 0 < t < 1,\ 0 < b < 1,$$
$$u(0) = u(1) = u'(0) = 0. \tag{5.3}$$

Since $\alpha = 5/2$, we have $C_1 = 0.1290$ and $C_2 = 0.0077$. Let $f(u) = u^b$, $0 < b < 1$. Then we have $F_\infty = 0$, $f_0 = +\infty$. Choose $l = 1/2$. Then $q(1/2) = \sqrt{2}/8 = 0.1768$. So $q(l)C_2 f_0 > F_\infty C_1$ holds. Thus, by Theorem 3.3, the boundary value problem (5.3) has a positive solution for each $\lambda \in (0, +\infty)$.

Example 5.3. Consider the boundary value problem

$$D_{0^+}^{5/2}u(t) + \lambda \frac{(200u^2 + u)(2 + \sin u)}{u + 1} = 0, \quad 0 < t < 1,\ a > 1,$$
$$u(0) = u(1) = u'(0) = 0. \tag{5.4}$$

Since $\alpha = 5/2$, we have $C_1 = 0.129$ and $C_2 = 0.0077$. Let $f(u) = (200u^2 + u)(2 + \sin u)/(u + 1)$. Then we have $F_0 = f_0 = 2$, $F_\infty = 600$, $f_\infty = 200$, and $2u < f(u) < 600u$.

(i) Choose $l = 1/2$. Then $q(1/2) = \sqrt{2}/8 = 0.1768$. So $q(l)C_2 f_\infty > F_0 C_1$ holds. Thus, by Theorem 3.2, the boundary value problem (5.4) has a positive solution for each $\lambda \in (3.6937, 3.8759)$.

(ii) By Theorem 4.1, the boundary value problem (5.4) has no positive solution for all $\lambda \in (0, 0.0129)$.

(iii) By Theorem 4.2, the boundary value problem (5.4) has no positive solution for all $\lambda \in (369.369, +\infty)$.

Example 5.4. Consider the boundary value problem

$$D_{0^+}^{5/2}u(t) + \lambda \frac{(u^2 + u)(2 + \sin u)}{150u + 1} = 0, \quad 0 < t < 1,\ a > 1,$$
$$u(0) = u(1) = u'(0) = 0. \tag{5.5}$$

Since $\alpha = 5/2$, we have $C_1 = 0.129$ and $C_2 = 0.0077$. Let $f(u) = (u^2 + u)(2 + \sin u)/(150u + 1)$. Then we have $F_0 = f_0 = 2$, $F_\infty = 1/50$, $f_\infty = 1/150$, and $u/150 < f(u) < 2u$.

(i) Choose $l = 1/2$. Then $q(1/2) = \sqrt{2}/8 = 0.1768$. So $q(l)C_2 f_0 > F_\infty C_1$ holds. Thus, by Theorem 3.3, the boundary value problem (5.5) has a positive solution for each $\lambda \in (369.369, 387.5968)$.

(ii) By Theorem 4.1, the boundary value problem (5.5) has no positive solution for all $\lambda \in (0, 3.8759)$.

(iii) By Theorem 4.2, the boundary value problem (5.5) has no positive solution for all $\lambda \in (110810.6911, +\infty)$.

to Contents, page 1

Acknowledgments

The authors sincerely thank the reviewers for their valuable suggestions and useful comments that have led to the present improved version of the original manuscript. This research is supported by the Natural Science Foundation of China (11071143, 11026112, 60904024), the Natural Science Foundation of Shandong (Y2008A28, ZR2009AL003), University of Jinan Research Funds for Doctors (XBS0843) and University of Jinan Innovation Funds for Graduate Students (YCX09014).

References

[1] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equation*, A Wiley-Interscience Publication, John Wiley & Sons, New York, NY, USA, 1993.
[2] K. B. Oldham and J. Spanier, *The Fractional Calculus*, Academic Press, New York, NY, USA, 1974.
[3] I. Podlubny, *Fractional Differential Equations*, vol. 198 of *Mathematics in Science and Engineering*, Academic Press, San Diego, Calif, USA, 1999.
[4] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional Integral and Derivative: Theory and Applications*, Gordon and Breach Science, Yverdon, Switzerland, 1993.
[5] Q. Li and S. Sun, "On the existence of positive solutions for initial value problem to a class of fractional differential equation," in *Proceedings of the 7th Conference on Biological Dynamic System and Stability of Differential Equation*, pp. 886–889, World Academic Press, Chongqing, China, 2010.
[6] Q. Li, S. Sun, M. Zhang, and Y. Zhao, "On the existence and uniqueness of solutions for initial value problem of fractional differential equations," *Journal of University of Jinan*, vol. 24, pp. 312–315, 2010.
[7] Q. Li, S. Sun, Z. Han, and Y. Zhao, "On the existence and uniqueness of solutions for initial value problem of nonlinear fractional differential equations," in *Proceedings of the 6th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA '10)*, pp. 452–457, Qingdao, China, 2010.
[8] D. Delbosco and L. Rodino, "Existence and uniqueness for a nonlinear fractional differential equation," *Journal of Mathematical Analysis and Applications*, vol. 204, no. 2, pp. 609–625, 1996.
[9] S. Zhang, "The existence of a positive solution for a nonlinear fractional differential equation," *Journal of Mathematical Analysis and Applications*, vol. 252, no. 2, pp. 804–812, 2000.
[10] S. Zhang, "Existence of positive solution for some class of nonlinear fractional differential equations," *Journal of Mathematical Analysis and Applications*, vol. 278, no. 1, pp. 136–148, 2003.
[11] H. Jafari and V. Daftardar-Gejji, "Positive solutions of nonlinear fractional boundary value problems using Adomian decomposition method," *Applied Mathematics and Computation*, vol. 180, no. 2, pp. 700–706, 2006.
[12] S. Zhang, "Positive solutions for boundary-value problems of nonlinear fractional differential equations," *Electronic Journal of Differential Equations*, vol. 36, pp. 1–12, 2006.
[13] T. Qiu and Z. Bai, "Existence of positive solutions for singular fractional differential equations," *Electronic Journal of Differential Equations*, vol. 146, pp. 1–9, 2008.
[14] Z. Bai and H. Lü, "Positive solutions for boundary value problem of nonlinear fractional differential equation," *Journal of Mathematical Analysis and Applications*, vol. 311, no. 2, pp. 495–505, 2005.
[15] M. Zhang, S. Sun, Y. Zhao, and D. Yang, "Existence of positive solutions for boundary value problems of fractional differential equations," *Journal of University of Jinan*, vol. 24, pp. 205–208, 2010.
[16] Y. Zhao and S. Sun, "On the existence of positive solutions for boundary value problems of nonlinear fractional differential equations," in *Proceedings of the 7th Conference on Biological Dynamic System and Stability of Differential Equation*, pp. 682–685, World Academic Press, Chongqing, China, 2010.
[17] Y. Zhao, S. Sun, Z. Han, and M. Zhang, "Existence on positive solutions for boundary value problems of singular nonlinear fractional differential equations," in *Proceedings of the 6th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications*, pp. 480–485, Qingdao, China, 2010.
[18] Y. Zhao, S. Sun, Z. Han, and Q. Li, "The existence of multiple positive solutions for boundary value problems of nonlinear fractional differential equations," *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, no. 4, pp. 2086–2097, 2011.
[19] Y. Yu and D. Jiang, "Multiple Positive Solutions for the Boundary Value Problem of A Nonlinear Fractional Differential Equation," Northeast Normal University, 2009.

to Contents, page 1

[20] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and Applications of Fractional Differential Equations*, vol. 204 of *North-Holland Mathematics Studies*, Elsevier Science, Amsterdam, The Netherlands, 2006.

[21] M. A. Krasnoselskii, *Positive Solutions of Operator Equations*, Noordhoff, Groningen, The Netherlands, 1964.

[22] W. Ge, *Boundary Value Problem of Nonlinear Ordinary Differential Equations*, Science Press, Beijing, China, 2007.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 857860, 15 pages
doi:10.1155/2011/857860

Research Article

Properties of Third-Order Nonlinear Functional Differential Equations with Mixed Arguments

B. Baculíková

Department of Mathematics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Letná 9, 042 00 Košice, Slovakia

Correspondence should be addressed to B. Baculíková, blanka.baculikova@tuke.sk

Received 14 December 2010; Accepted 20 January 2011

Academic Editor: Josef Diblík

The aim of this paper is to offer sufficient conditions for property (B) and/or the oscillation of the third-order nonlinear functional differential equation with mixed arguments $[a(t)[x''(t)]^{\gamma}]' = q(t)f(x[\tau(t)])+p(t)h(x[\sigma(t)])$. Both cases $\int^{\infty} a^{-1/\gamma}(s)\mathrm{d}s = \infty$ and $\int^{\infty} a^{-1/\gamma}(s)\mathrm{d}s < \infty$ are considered. We deduce properties of the studied equations via new comparison theorems. The results obtained essentially improve and complement earlier ones.

1. Introduction

We are concerned with the oscillatory and certain asymptotic behavior of all solutions of the third-order functional differential equations

$$\left[a(t)\left[x''(t)\right]^{\gamma}\right]' = q(t)f(x[\tau(t)]) + p(t)h(x[\sigma(t)]). \tag{E}$$

Throughout the paper, it is assumed that $a, q, p \in C([t_0, \infty))$, $\tau, \sigma \in C^1([t_0, \infty))$, $f, h \in C((-\infty, \infty))$, and

(H_1) γ is the ratio of two positive odd integers,

(H_2) $a(t)$, $q(t)$, $p(t)$ are positive,

(H_3) $\tau(t) \le t$, $\sigma(t) \ge t$, $\tau'(t) > 0$, $\sigma'(t) > 0$, $\lim_{t\to\infty}\tau(t) = \infty$,

(H_4) $f^{1/\gamma}(x)/x \ge 1$, $xh(x) > 0$, $f'(x) \ge 0$, and $h'(x) \ge 0$ for $x \ne 0$,

(H_5) $-f(-xy) \ge f(xy) \ge f(x)f(y)$ for $xy > 0$ and $-h(-xy) \ge h(xy) \ge h(x)h(y)$ for $xy > 0$.

to Contents, page 1

By a solution of (E), we mean a function $x(t) \in C^2([T_x,\infty))$, $T_x \geq t_0$, which has the property $a(t)(x''(t))^\gamma \in C^1([T_x,\infty))$ and satisfies (E) on $[T_x,\infty)$. We consider only those solutions $x(t)$ of (E) which satisfy $\sup\{|x(t)| : t \geq T\} > 0$ for all $T \geq T_x$. We assume that (E) possesses such a solution. A solution of (E) is called oscillatory if it has arbitrarily large zeros on $[T_x,\infty)$, and, otherwise, it is nonoscillatory. Equation (E) is said to be oscillatory if all its solutions are oscillatory.

Recently, (E) and its particular cases (see [1–17]) have been intensively studied. The effort has been oriented to provide sufficient conditions for every (E) to satisfy

$$\lim_{t\to\infty}|x(t)| = \infty \tag{1.1}$$

or to eliminate all nonoscillatory solutions. Following [6, 8, 13, 15], we say that (E) has property (B) if each of its nonoscillatory solutions satisfies (1.1).

We will discuss both cases

$$\int_{t_0}^{\infty} a^{-1/\gamma}(s)\mathrm{d}s < \infty, \tag{1.2}$$

$$\int_{t_0}^{\infty} a^{-1/\gamma}(s)\mathrm{d}s = \infty. \tag{1.3}$$

We will establish suitable comparison theorems that enable us to study properties of (E) regardless of the fact that (1.3) or (1.2) holds. We will compare (E) with the first-order advanced/delay equations, in the sense that the oscillation of these first-order equations yields property (B) or the oscillation of (E).

In the paper, we are motivated by an interesting result of Grace et al. [10], where the oscillation criteria for (E) are discussed. This result has been complemented by Baculíková et al. [5]. When studying properties of (E), the authors usually reduce (E) onto the corresponding differential inequalities

$$\begin{aligned}
&\left[a(t)\left[x''(t)\right]^\gamma\right]' \geq q(t)f(x[\tau(t)]),\\
&\left[a(t)\left[x''(t)\right]^\gamma\right]' \geq p(t)h(x[\sigma(t)]),
\end{aligned} \tag{E_σ}$$

and further study only properties of these inequalities. Therefore, the criteria obtained withhold information either from delay argument $\tau(t)$ and the corresponding functions $q(t)$ and $f(u)$ or from advanced argument $\sigma(t)$ and the corresponding functions $p(t)$ and $h(u)$. In the paper, we offer a technique for obtaining new criteria for property (B) and the oscillation of (E) that involve both arguments $\tau(t)$ and $\sigma(t)$. Consequently, our results are new even for the linear case of (E) and properly complement and extend earlier ones presented in [1–17].

Remark 1.1. All functional inequalities considered in this paper are assumed to hold eventually; that is, they are satisfied for all t large enough.

2. Main Results

The following results are elementary but useful in what comes next.

to Contents, page 1

Lemma 2.1. *Assume that* $A \geq 0$, $B \geq 0$, $\alpha \geq 1$. *Then,*

$$(A+B)^{\alpha} \geq A^{\alpha} + B^{\alpha}. \tag{2.1}$$

Proof. If $A = 0$ or $B = 0$, then (2.1) holds. For $A \neq 0$, setting $x = B/A$, condition (2.1) takes the form $(1+x)^{\alpha} \geq 1 + x^{\alpha}$, which is for $x > 0$ evidently true. □

Lemma 2.2. *Assume that* $A \geq 0$, $B \geq 0$, $0 < \alpha \leq 1$. *Then,*

$$(A+B)^{\alpha} \geq \frac{A^{\alpha} + B^{\alpha}}{2^{1-\alpha}}. \tag{2.2}$$

Proof. We may assume that $0 < A < B$. Consider a function $g(u) = u^{\alpha}$. Since $g''(u) < 0$ for $u > 0$, function $g(u)$ is concave down; that is,

$$g\left(\frac{A+B}{2}\right) \geq \frac{g(A) + g(B)}{2} \tag{2.3}$$

which implies (2.2). □

The following result presents a useful relationship between an existence of positive solutions of the advanced differential inequality and the corresponding advanced differential equation.

Lemma 2.3. *Suppose that* $p(t)$, $\sigma(t)$, *and* $h(u)$ *satisfy* (H_2), (H_3), *and* (H_4), *respectively. If the first-order advanced differential inequality*

$$z'(t) - p(t)h(z(\sigma(t))) \geq 0 \tag{2.4}$$

has an eventually positive solution, so does the advanced differential equation

$$z'(t) - p(t)h(z(\sigma(t))) = 0. \tag{2.5}$$

Proof. Let $z(t)$ be a positive solution of (2.4) on $[t_1, \infty)$. Then, $z(t)$ satisfies the inequality

$$z(t) \geq z(t_1) + \int_{t_1}^{t} p(s)h(z(\sigma(s)))\mathrm{d}s. \tag{2.6}$$

Let

$$\begin{aligned} y_1(t) &= z(t), \\ y_n(t) &= z(t_1) + \int_{t_1}^{t} p(s)h\big(y_{n-1}(\sigma(s))\big)\mathrm{d}s, \qquad n = 2, 3 \ldots. \end{aligned} \tag{2.7}$$

to Contents, page 1

It follows from the definition of $y_n(t)$ and (H_4) that the sequence $\{y_n\}$ has the property

$$z(t) = y_1(t) \geq y_2(t) \geq \cdots \geq z(t_1), \qquad t \geq t_1. \tag{2.8}$$

Hence, $\{y_n\}$ converges pointwise to a function $y(t)$, where $z(t) \geq y(t) \geq z(t_1)$. Let $h_n(t) = p(t)h(y_n(\sigma(t)))$, $n = 1, 2, \ldots$, then $h_1(t) \geq h_2(t) \geq \cdots \geq 0$. Since $h_1(t)$ is integrable on $[t_1, t]$ and $\lim_{n\to\infty} h_n(t) = p(t)h(y(\sigma(t)))$, it follows by Lebesgue's dominated convergence theorem that

$$y(t) = z(t_1) + \int_{t_1}^{t} p(s)h\big(y(\sigma(s))\big)\mathrm{d}s. \tag{2.9}$$

Thus, $y(t)$ satisfies (2.5). $\square$

We start our main results with the classification of the possible nonoscillatory solutions of (E).

Lemma 2.4. *Let $x(t)$ be a nonoscillatory solution of (E). Then, $x(t)$ satisfies, eventually, one of the following conditions*

(I)

$$x(t)x'(t) > 0, \qquad x(t)x''(t) > 0, \qquad x(t)\left[a(t)\left[x''(t)\right]^{\gamma}\right]' > 0, \tag{2.10}$$

(II)

$$x(t)x'(t) > 0, \qquad x(t)x''(t) < 0, \qquad x(t)\left[a(t)\left[x''(t)\right]^{\gamma}\right]' > 0, \tag{2.11}$$

and if (1.2) holds, then also

(III)

$$x(t)x'(t) < 0, \qquad x(t)x''(t) > 0, \qquad x(t)\left[a(t)\left[x''(t)\right]^{\gamma}\right]' > 0. \tag{2.12}$$

Proof. Let $x(t)$ be a nonoscillatory solution of (E), say $x(t) > 0$ for $t \geq t_0$. It follows from (E) that $[a(t)[x''(t)]^{\gamma}] > 0$, eventually. Thus, $x''(t)$ and $x'(t)$ are of fixed sign for $t \geq t_1$, t_1 large enough. At first, we assume that $x''(t) < 0$. Then, either $x'(t) > 0$ or $x'(t) < 0$, eventually. But $x''(t) < 0$ together with $x'(t) < 0$ imply that $x(t) < 0$. A contradiction, that is, Case (II) holds.

Now, we suppose that $x''(t) > 0$, then either Case (I) or Case (III) holds. On the other hand, if (1.3) holds, then Case (III) implies that $a(t)[x''(t)]^{\gamma} \geq c > 0$, $t \geq t_1$. Integrating from t_1 to t, we have

$$x'(t) - x'(t_1) \geq c^{1/\gamma} \int_{t_1}^{t} a^{-1/\gamma}(s)\mathrm{d}s, \tag{2.13}$$

which implies that $x'(t) \to \infty$ as $t \to \infty$, and we deduce that Case (III) may occur only if (1.2) is satisfied. The proof is complete. $\square$

to Contents, page 1

Remark 2.5. It follows from Lemma 2.4 that if (1.3) holds, then only Cases (I) and (II) may occur.

In the following results, we provide criteria for the elimination of Cases (I)–(III) of Lemma 2.4 to obtain property (B)/oscillation of (E).

Let us denote for our further references that

$$P(t) = \int_t^\infty a^{-1/\gamma}(u)\left(\int_u^\infty p(s)\mathrm{d}s\right)^{1/\gamma}\mathrm{d}u, \tag{2.14}$$

$$Q(t) = \int_t^\infty a^{-1/\gamma}(u)\left(\int_u^\infty \frac{q(\tau^{-1}(s))}{\tau'(\tau^{-1}(s))}\mathrm{d}s\right)^{1/\gamma}\mathrm{d}u. \tag{2.15}$$

Theorem 2.6. *Let $0 < \gamma \le 1$. Assume that $x(t)$ is a nonoscillatory solution of (E). If the first-order advanced differential equation*

$$z'(t) - P(t)e^{-\int_{t_1}^t Q(s)\mathrm{d}s}\, h^{1/\gamma}\left(e^{\int_{t_1}^{\sigma(t)} Q(s)\mathrm{d}s}\right)h^{1/\gamma}(z[\sigma(t)]) = 0 \tag{E_1}$$

is oscillatory, then Case (II) cannot hold.

Proof. Let $x(t)$ be a nonoscillatory solution of (E), satisfying Case (II) of Lemma 2.4. We may assume that $x(t) > 0$ for $t \ge t_0$. Integrating (E) from t to ∞, one gets

$$-a(t)[x''(t)]^\gamma \ge \int_t^\infty q(s)f(x[\tau(s)])\mathrm{d}s + \int_t^\infty p(s)h(x[\sigma(s)])\mathrm{d}s. \tag{2.16}$$

On the other hand, the substitution $\tau(s) = u$ gives

$$\begin{aligned}\int_t^\infty q(s)f(x[\tau(s)])\mathrm{d}s &= \int_{\tau(t)}^\infty \frac{q(\tau^{-1}(u))}{\tau'(\tau^{-1}(u))}f(x(u))\mathrm{d}u \\ &\ge \int_t^\infty \frac{q(\tau^{-1}(s))}{\tau'(\tau^{-1}(s))}f(x(s))\mathrm{d}s.\end{aligned} \tag{2.17}$$

Using (2.17) in (2.16), we find

$$-x''(t) \ge a^{-1/\gamma}(t)\left(\int_t^\infty \frac{q(\tau^{-1}(s))}{\tau'(\tau^{-1}(s))}f(x(s))\mathrm{d}s + \int_t^\infty p(s)h(x[\sigma(s)])\mathrm{d}s\right)^{1/\gamma}. \tag{2.18}$$

to Contents, page 1

Taking into account the monotonicity of $x(t)$, it follows from Lemma 2.1 that

$$\begin{aligned}-x''(t) \geq{}& \frac{f^{1/\gamma}(x(t))}{a^{1/\gamma}(t)}\left(\int_t^\infty \frac{q(\tau^{-1}(s))}{\tau'(\tau^{-1}(s))}\mathrm{d}s\right)^{1/\gamma}\\ &+\frac{h^{1/\gamma}(x[\sigma(t)])}{a^{1/\gamma}(t)}\left(\int_t^\infty p(s)\mathrm{d}s\right)^{1/\gamma},\end{aligned} \tag{2.19}$$

where we have used (H_3) and (H_4). An integration from t to ∞ yields

$$\begin{aligned}x'(t) \geq{}& \int_t^\infty \frac{f^{1/\gamma}(x(u))}{a^{1/\gamma}(u)}\left(\int_u^\infty \frac{q(\tau^{-1}(s))}{\tau'(\tau^{-1}(s))}\mathrm{d}s\right)^{1/\gamma}\mathrm{d}u\\ &+\int_t^\infty \frac{h^{1/\gamma}(x[\sigma(u)])}{a^{1/\gamma}(u)}\left(\int_u^\infty p(s)\mathrm{d}s\right)^{1/\gamma}\mathrm{d}u\\ \geq{}& f^{1/\gamma}(x(t))Q(t)+h^{1/\gamma}(x[\sigma(t)])P(t).\end{aligned} \tag{2.20}$$

Regarding (H_4), it follows that $x(t)$ is a positive solution of the differential inequality

$$x'(t)-Q(t)x(t) \geq P(t)h^{1/\gamma}(x[\sigma(t)]). \tag{2.21}$$

Applying the transformation

$$x(t)=w(t)e^{\int_{t_1}^t Q(s)\mathrm{d}s}, \tag{2.22}$$

we can easily verify that $w(t)$ is a positive solution of the advanced differential inequality

$$w'(t)-P(t)e^{-\int_{t_1}^t Q(s)\mathrm{d}s}\, h^{1/\gamma}\left(e^{\int_{t_1}^{\sigma(t)} Q(s)\mathrm{d}s}\right)h^{1/\gamma}(w[\sigma(t)]) \geq 0. \tag{2.23}$$

By Lemma 2.3, we conclude that the corresponding differential equation (E_1) has also a positive solution. A contradiction. Therefore, $x(t)$ cannot satisfy Case (II). □

to Contents, page 1

Remark 2.7. It follows from the proof of Theorem 2.8 that if at least one of the following conditions is satisfied:

$$\begin{gathered}
\int_{t_0}^{\infty} p(s)\mathrm{d}s = \infty, \\
\int_{t_0}^{\infty} \frac{q(\tau^{-1}(s))}{\tau'(\tau^{-1}(s))}\mathrm{d}s = \infty, \\
\int_{t_0}^{\infty} a^{-1/\gamma}(u)\left(\int_u^{\infty} p(s)\mathrm{d}s\right)^{1/\gamma}\mathrm{d}u = \infty, \\
\int_{t_0}^{\infty} a^{-1/\gamma}(u)\left(\int_u^{\infty} \frac{q(\tau^{-1}(s))}{\tau'(\tau^{-1}(s))}\mathrm{d}s\right)^{1/\gamma}\mathrm{d}u = \infty,
\end{gathered} \tag{2.24}$$

then any nonoscillatory solution $x(t)$ of (E) cannot satisfy Case (II). Therefore, we may assume that the corresponding integrals in (2.14)-(2.15) are convergent.

Now, we are prepared to provide new criteria for property (B) of (E) and also the rate of divergence of all nonoscillatory solutions.

Theorem 2.8. *Let* (1.3) *hold and* $0 < \gamma \le 1$. *Assume that* (E_1) *is oscillatory. Then,* (E) *has property (B) and, what is more, the following rate of divergence for each of its nonoscillatory solutions holds:*

$$|x(t)| \ge c\int_{t_1}^{t} a^{-1/\gamma}(s)(t-s)\mathrm{d}s, \qquad c > 0. \tag{2.25}$$

Proof. Let $x(t)$ be a positive solution of (E). It follows from Lemma 2.4 and Remark 2.5 that $x(t)$ satisfies either Case (I) or (II). But Theorem 2.6 implies that the Case (II) cannot hold. Therefore, $x(t)$ satisfies Case (I), which implies (1.1); that is, (E) has property (B). On the other hand, there is a constant $c > 0$ such that

$$a(t)\left(x''(t)\right)^{\gamma} \ge c^{\gamma}. \tag{2.26}$$

Integrating twice from t_1 to t, we have

$$x(t) \ge c\int_{t_1}^{t}\left(\int_{t_1}^{u} a^{-1/\gamma}(s)\mathrm{d}s\right)\mathrm{d}u = c\int_{t_1}^{t} a^{-1/\gamma}(s)(t-s)\mathrm{d}s, \tag{2.27}$$

which is the desired estimate. □

Employing an additional condition on the function $h(x)$, we get easily verifiable criterion for property (B) of (E).

to Contents, page 1

Corollary 2.9. *Let* $0 < \gamma \leq 1$ *and* (1.3) *hold. Assume that*

$$h^{1/\gamma}(x)/x \geq 1, \qquad |x| \geq 1, \tag{2.28}$$

$$\liminf_{t\to\infty} \int_t^{\sigma(t)} P(u) e^{\int_u^{\sigma(u)} Q(s)\mathrm{d}s} \mathrm{d}u > \frac{1}{e}. \tag{2.29}$$

Then, (E) *has property (B).*

Proof. First note that (2.29) implies

$$\int_{t_0}^{\infty} P(u) e^{\int_u^{\sigma(u)} Q(s)\mathrm{d}s} \mathrm{d}u = \infty. \tag{2.30}$$

By Theorem 2.8, it is sufficient to show that (E_1) is oscillatory. Assume the converse, let (E_1) have an eventually positive solution $z(t)$. Then, $z'(t) > 0$ and so $z(\sigma(t)) > c > 0$. Integrating (E_1) from t_1 to t, we have in view of (2.28)

$$\begin{aligned} z(t) &\geq \int_{t_1}^{t} P(u) e^{-\int_{t_1}^{u} Q(s)\mathrm{d}s} h^{1/\gamma}\left(e^{\int_{t_1}^{\sigma(u)} Q(s)\mathrm{d}s}\right) h^{1/\gamma}(z[\sigma(u)])\mathrm{d}u \\ &\geq h^{1/\gamma}(c) \int_{t_1}^{t} P(u) e^{\int_u^{\sigma(u)} Q(s)\mathrm{d}s} \mathrm{d}u. \end{aligned} \tag{2.31}$$

Using (2.30) in the previous inequalities, we get $z(t) \to \infty$ as $t \to \infty$. Therefore, $z(t) \geq 1$, eventually. Now, using (2.28) in (E_1), one can verify that $z(t)$ is a positive solution of the differential inequality

$$z'(t) - P(t) e^{\int_t^{\sigma(t)} Q(s)\mathrm{d}s} z(\sigma(t)) \geq 0. \tag{2.32}$$

But, by [14, Theorem 2.4.1], condition (2.29) ensures that (2.32) has no positive solutions. This is a contradiction, and we conclude that (E) has property (B). □

Example 2.10. Consider the third-order nonlinear differential equation with mixed arguments

$$\left(t^{1/3}\left(x''(t)\right)^{1/3}\right)' = \frac{a}{t^{4/3}} x^{1/3}(\lambda t) + \frac{b}{t^{4/3}} x^{\beta}(\omega t), \tag{E_{x1}}$$

where $a, b > 0$, $0 < \lambda < 1$, $\omega > 1$, and $\beta \geq 1/3$ is a ratio of two positive odd integers. Since

$$P(t) = \frac{27b^3}{t}, \qquad Q(t) = \frac{27a^3\lambda}{t}, \tag{2.33}$$

Corollary 2.9 implies that (E_{x1}) has property (B) provided that

$$b^3 \omega^{27a^3\lambda} \ln \omega > \frac{1}{27\mathrm{e}}. \tag{2.34}$$

to Contents, page 1

Moreover, by Theorem 2.8, the rate of divergence of every nonoscillatory solution of (E_{x1}) is

$$|x(t)| \geq ct \ln t, \qquad c > 0. \tag{2.35}$$

For $\beta = 1/3$ and $\delta > 1$ satisfying $\delta^{1/3}(\delta-1)^{4/3} = 3a\lambda^{\delta/3} + 3b\omega^{\delta/3}$, one such solution is t^{δ}.

Now, we turn our attention to the case when $\gamma \geq 1$.

Theorem 2.11. *Let $\gamma \geq 1$. Assume that $x(t)$ is a nonoscillatory solution of (E). If the first-order advanced differential equation*

$$z'(t) - 2^{(1-\gamma)/\gamma} P(t) e^{[-2^{(1-\gamma)/\gamma} \int_{t_1}^{t} Q(s)\mathrm{d}s]} h^{1/\gamma}\left(e^{2^{(1-\gamma)/\gamma} \int_{t_1}^{\sigma(t)} Q(s)\mathrm{d}s}\right) h^{1/\gamma}(z[\sigma(t)]) = 0 \tag{E_2}$$

is oscillatory, then Case (II) cannot hold.

Proof. Let $x(t)$ be an eventually positive solution of (E), satisfying Case (II) of Lemma 2.4. Then, (2.18) holds. Lemma 2.2, in view of the monotonicity of $x(t)$, (H_3), and (H_4), implies

$$\begin{aligned} -x''(t) \geq{}& \frac{f^{1/\gamma}(x(t))}{2^{(\gamma-1)/\gamma} a^{1/\gamma}(t)} \left(\int_t^{\infty} \frac{q(\tau^{-1}(s))}{\tau'(\tau^{-1}(s))} \mathrm{d}s \right)^{1/\gamma} \\ &+ \frac{h^{1/\gamma}(x[\sigma(t)])}{2^{(\gamma-1)/\gamma} a^{1/\gamma}(t)} \left(\int_t^{\infty} p(s)\mathrm{d}s \right)^{1/\gamma}. \end{aligned} \tag{2.36}$$

An integration from t to ∞ yields

$$\begin{aligned} x'(t) \geq{}& \int_t^{\infty} \frac{f^{1/\gamma}(x(u))}{2^{(\gamma-1)/\gamma} a^{1/\gamma}(u)} \left(\int_u^{\infty} \frac{q(\tau^{-1}(s))}{\tau'(\tau^{-1}(s))} \mathrm{d}s \right)^{1/\gamma} \mathrm{d}u \\ &+ \int_t^{\infty} \frac{h^{1/\gamma}(x[\sigma(u)])}{2^{(\gamma-1)/\gamma} a^{1/\gamma}(u)} \left(\int_u^{\infty} p(s)\mathrm{d}s \right)^{1/\gamma} \mathrm{d}u \\ \geq{}& f^{1/\gamma}(x(t)) 2^{(1-\gamma)/\gamma} Q(t) + h^{1/\gamma}(x[\sigma(t)]) 2^{(1-\gamma)/\gamma} P(t). \end{aligned} \tag{2.37}$$

Noting (H_4), we see that $x(t)$ is a positive solution of the differential inequality

$$x'(t) \geq 2^{(1-\gamma)/\gamma} Q(t) x(t) + 2^{(1-\gamma)/\gamma} P(t) h^{1/\gamma}(x[\sigma(t)]). \tag{2.38}$$

Setting

$$x(t) = w(t) e^{[2^{(1-\gamma)/\gamma} \int_{t_1}^{t} Q(s)\mathrm{d}s]}, \tag{2.39}$$

to Contents, page 1

one can see that $w(t)$ is a positive solution of the advanced differential inequality

$$w'(t) - 2^{(1-\gamma)/\gamma} P(t) e^{[-2^{(1-\gamma)/\gamma} \int_{t_1}^{t} Q(s)\mathrm{d}s]} \, h^{1/\gamma}\left(e^{2^{(1-\gamma)/\gamma} \int_{t_1}^{\sigma(t)} Q(s)\mathrm{d}s}\right) h^{1/\gamma}(w[\sigma(t)]) \geq 0. \tag{2.40}$$

By Lemma 2.3, we deduce that the corresponding differential equation (E_2) has also a positive solution. A contradiction. Therefore, $x(t)$ cannot satisfy Case (II). □

The following result is obvious.

Theorem 2.12. *Let* (1.3) *hold and* $\gamma \geq 1$. *Assume that* (E_2) *is oscillatory. Then,* (E) *has property (B) and, what is more, each of its nonoscillatory solutions satisfies* (2.25).

Now, we present easily verifiable criterion for property (B) of (E).

Corollary 2.13. *Let* (1.3) *and* (2.28) *hold and* $\gamma \geq 1$. *If*

$$\liminf_{t \to \infty} \int_{t}^{\sigma(t)} P(u) e^{[2^{(1-\gamma)/\gamma} \int_{u}^{\sigma(u)} Q(s)\mathrm{d}s]} \mathrm{d}u > \frac{2^{(\gamma-1)/\gamma}}{e}, \tag{2.41}$$

then (E) *has property (B).*

Proof. The proof is similar to the proof of Corollary 2.9 and so it can be omitted. □

Remark 2.14. Theorems 2.6, 2.8, 2.11, and 2.12 and Corollaries 2.9 and 2.13 provide criteria for property (B) that include both delay and advanced arguments and all coefficients and functions of (E). Our results are new even for the linear case of (E).

Remark 2.15. It is useful to notice that if we apply the traditional approach to (E), that is, if we replace (E) by the corresponding differential inequality (E_σ), then conditions (2.29) of Corollary 2.9 and (2.41) of Corollary 2.13 would take the forms

$$\liminf_{t \to \infty} \int_{t}^{\sigma(t)} P(u)\mathrm{d}u > \frac{1}{e}, \qquad \liminf_{t \to \infty} \int_{t}^{\sigma(t)} P(u)\mathrm{d}u > \frac{2^{(\gamma-1)/\gamma}}{e}, \tag{2.42}$$

respectively, which are evidently second to (2.29) and (2.41).

Example 2.16. Consider the third-order nonlinear differential equation with mixed arguments

$$\left(t(x''(t))^3\right)' = \frac{a}{t^6} x^3(\lambda t) + \frac{b}{t^6} x^\beta(\omega t), \tag{E_{x2}}$$

where $a, b > 0$, $0 < \lambda < 1$, $\beta \geq 3$ is a ratio of two positive odd integers and $\omega > 1$. It is easy to see that conditions (2.14) and (2.15) for (E_{x2}) reduce to

$$P(t) = \frac{b^{1/3}}{5^{1/3} t}, \qquad Q(t) = \frac{\lambda^{5/3} a^{1/3}}{5^{1/3} t}, \tag{2.43}$$

to Contents, page 1

respectively. It follows from Corollary 2.13 that (E_{x2}) has property (B) provided that

$$b^{1/3}\left[\omega^{\lambda^{5/3}a^{1/3}/2^{2/3}5^{1/3}}\right]\ln\omega \geq \frac{2^{2/3}5^{1/3}}{e}. \tag{2.44}$$

Moreover, (2.25) provides the following rate of divergence for every nonoscillatory solution of (E_{x2}):

$$|x(t)| \geq ct^{5/3}, \qquad c > 0. \tag{2.45}$$

Now, we eliminate Case (I) of Lemma 2.4, to get the oscillation of (E).

Theorem 2.17. *Let $x(t)$ be a nonoscillatory solution of (E). Assume that there exists a function $\xi(t) \in C^1([t_0,\infty))$ such that*

$$\xi'(t) \geq 0, \quad \xi(t) < t, \qquad \eta(t) = \sigma(\xi(\xi(t))) > t. \tag{2.46}$$

If the first-order advanced differential equation

$$z'(t) - \left\{\int_{\xi(t)}^{t} a^{-1/\gamma}(u)\left(\int_{\xi(u)}^{u} p(s)\mathrm{d}s\right)^{1/\gamma}\mathrm{d}u\right\}h^{1/\gamma}(z[\eta(t)]) = 0 \tag{E_3}$$

is oscillatory, then Case (I) cannot hold.

Proof. Let $x(t)$ be an eventually positive solution of (E), satisfying Case (I). It follows from (E) that

$$\left[a(t)[x''(t)]^{\gamma}\right]' \geq p(t)h(x[\sigma(t)]). \tag{2.47}$$

Integrating from $\xi(t)$ to t, we have

$$\begin{aligned} a(t)[x''(t)]^{\gamma} - a(\xi(t))[x''(\xi(t))]^{\gamma} &\geq \int_{\xi(t)}^{t} p(s)h(x[\sigma(s)])\mathrm{d}s \\ &\geq h(x[\sigma(\xi(t))])\int_{\xi(t)}^{t} p(s)\mathrm{d}s. \end{aligned} \tag{2.48}$$

Therefore,

$$x''(t) \geq h^{1/\gamma}(x[\sigma(\xi(t))])a^{-1/\gamma}(t)\left(\int_{\xi(t)}^{t} p(s)\mathrm{d}s\right)^{1/\gamma}. \tag{2.49}$$

to Contents, page 1

An integration from $\xi(t)$ to t yields

$$\begin{aligned} x'(t) &\geq \int_{\xi(t)}^{t} h^{1/\gamma}(x[\sigma(\xi(u))])a^{-1/\gamma}(u)\left(\int_{\xi(u)}^{u} p(s)\mathrm{d}s\right)^{1/\gamma}\mathrm{d}u \\ &\geq h^{1/\gamma}(x[\eta(t)])\int_{\xi(t)}^{t} a^{-1/\gamma}(u)\left(\int_{\xi(u)}^{u} p(s)\mathrm{d}s\right)^{1/\gamma}\mathrm{d}u. \end{aligned} \tag{2.50}$$

Consequently, $x(t)$ is a positive solution of the advanced differential inequality

$$x'(t) - \left\{\int_{\xi(t)}^{t} a^{-1/\gamma}(u)\left(\int_{\xi(u)}^{u} p(s)\mathrm{d}s\right)^{1/\gamma}\mathrm{d}u\right\}h^{1/\gamma}(x[\eta(t)]) \geq 0. \tag{2.51}$$

Hence, by Lemma 2.3, we conclude that the corresponding differential equation (E_3) also has a positive solution, which contradicts the oscillation of (E_3). Therefore, $x(t)$ cannot satisfy Case (I). □

Combining Theorem 2.17 with Theorems 2.6 and 2.11, we get two criteria for the oscillation of (E).

Theorem 2.18. *Let* (1.3) *hold and* $0 < \gamma \leq 1$. *Assume that both of the first-order advanced equations* (E_1) *and* (E_3) *are oscillatory, then* (E) *is oscillatory.*

Proof. Assume that (E) has a nonoscillatory solution. It follows from Remark 2.5 that $x(t)$ satisfies either Case (I) or (II). But both cases are excluded by the oscillation of (E_1) and (E_3). □

Corollary 2.19. *Let* $0 < \gamma \leq 1$. *Assume that* (1.3), (2.28), (2.29), *and* (2.46) *hold. If*

$$\liminf_{t\to\infty}\int_{t}^{\eta(t)}\left\{\int_{\xi(v)}^{v} a^{-1/\gamma}(u)\left(\int_{\xi(u)}^{u} p(s)\mathrm{d}s\right)^{1/\gamma}\mathrm{d}u\right\}\mathrm{d}v > \frac{1}{e}, \tag{2.52}$$

then (E) *is oscillatory.*

Proof. Conditions (2.29) and (2.52) guarantee the oscillation of (E_1) and (E_3), respectively. The assertion now follows from Theorem 2.18. □

Example 2.20. We consider once more the third-order differential equation (E_{x1}) with the same restrictions as in Example 2.10. We set $\xi(t) = \alpha_0 t$, where $\alpha_0 = (1+\sqrt{\omega})/2\sqrt{\omega}$. Then condition (2.52) takes the form

$$b^3\frac{(1-\alpha_0)\left(1-\alpha_0^{1/3}\right)^3}{\alpha_0^2}\ln\left(\omega\alpha_0^2\right) > \frac{1}{27e}, \tag{2.53}$$

which by Corollary 2.19, implies the oscillation of (E_{x1}).

to Contents, page 1

The following results are obvious.

Theorem 2.21. *Let* (1.3) *hold and* $\gamma \geq 1$. *Assume that both of the first-order advanced equations* (E_2) *and* (E_3) *are oscillatory, then* (E) *is oscillatory.*

Corollary 2.22. *Let* $\gamma \geq 1$. *Assume that* (1.3), (2.28), (2.41), (2.46), *and* (2.52) *hold. Then* (E) *is oscillatory.*

Example 2.23. We recall again the differential equation (E_{x2}) with the same assumptions as in Example 2.16. We set $\xi(t) = \alpha_0 t$ with $\alpha_0 = (1+\sqrt{\omega})/2\sqrt{\omega}$. Then condition (2.52) reduces to

$$b^{1/3}\frac{(1-\alpha_0)\left(1-\alpha_0^5\right)^{1/3}}{\alpha_0^{8/3}}\ln\left(\omega\alpha_0^2\right) > \frac{5^{1/3}}{e}, \tag{2.54}$$

which, by Corollary 2.22, guarantees the oscillation of (E_{x2}).

The following result is intended to exclude Case (III) of Lemma 2.4.

Theorem 2.24. *Let* $x(t)$ *be a nonoscillatory solution of* (E). *Assume that* (1.2) *holds. If the first-order delay differential equation*

$$z'(t) + \left(\int_{t_1}^{t} q(s)\mathrm{d}s\right)^{1/\gamma}\left(\int_{t}^{\infty} a^{-1/\gamma}(s)\mathrm{d}s\right) f^{1/\gamma}(z[\tau(t)]) = 0. \tag{E_4}$$

is oscillatory, then Case (III) cannot hold.

Proof. Let $x(t)$ be a positive solution of (E), satisfying Case (III) of Lemma 2.4. Using that $a(t)[x''(t)]^{\gamma}$ is increasing, we find that

$$\begin{aligned} -x'(t) &\geq \int_{t}^{\infty} x''(s)\mathrm{d}s = \int_{t}^{\infty}\left(a^{1/\gamma}(s)x''(s)\right)a^{-1/\gamma}(s)\mathrm{d}s \\ &\geq a(t)^{1/\gamma}x''(t)\int_{t}^{\infty} a^{-1/\gamma}(s)\mathrm{d}s. \end{aligned} \tag{2.55}$$

Integrating the inequality $[a(t)[x''(t)]^{\gamma}]' \geq q(t)f(x[\tau(t)])$ from t_1 to t, we have

$$a(t)\left[x''(t)\right]^{\gamma} \geq \int_{t_1}^{t} q(s)f(x[\tau(s)]\mathrm{d}s) \geq f(x[\tau(t)])\int_{t_1}^{t} q(s)\mathrm{d}s. \tag{2.56}$$

Thus,

$$a^{1/\gamma}(t)x''(t) \geq f^{1/\gamma}(x[\tau(t)])\left(\int_{t_1}^{t} q(s)\mathrm{d}s\right)^{1/\gamma}. \tag{2.57}$$

to Contents, page 1

Combining (2.57) with (2.55), we find

$$0 \geq x'(t) + \left(\int_{t_1}^{t} q(s)\mathrm{d}s\right)^{1/\gamma} \left(\int_{t}^{\infty} a^{-1/\gamma}(s)\mathrm{d}s\right) f^{1/\gamma}(x[\tau(t)]). \tag{2.58}$$

It follows from [16, Theorem 1] that the corresponding differential equation (E_4) also has a positive solution. A contradiction. For that reason, $x(t)$ cannot satisfy Case (III). □

The following results are immediate.

Theorem 2.25. *Let* (1.2) *hold and* $0 < \gamma \leq 1$. *Assume that both of the first-order advanced equations* (E_1) *and* (E_4) *are oscillatory, then* (E) *has property (B).*

Theorem 2.26. *Let* (1.2) *hold and* $0 < \gamma \leq 1$. *Assume that all of the three first-order advanced equations* (E_1), (E_3), *and* (E_4) *are oscillatory, then* (E) *is oscillatory.*

Theorem 2.27. *Let* (1.2) *hold and* $\gamma \geq 1$. *Assume that both of the first-order advanced equations* (E_2) *and* (E_4) *are oscillatory, then* (E) *has property (B).*

Theorem 2.28. *Let* (1.2) *hold and* $\gamma \geq 1$. *Assume that all of the three first-order advanced equations* (E_2), (E_3), *and* (E_4) *are oscillatory, then* (E) *is oscillatory.*

3. Summary

In this paper, we have presented new comparison theorems for deducing the property (B)/oscillation of (E) from the oscillation of a set of the suitable first-order delay/advanced differential equation. We were able to present such criteria for studied properties that employ all coefficients and functions included in studied equations. Our method essentially simplifies the examination of the third-order equations, and, what is more, it supports backward the research on the first-order delay/advanced differential equations. Our results here extend and complement latest ones of Grace et al. [10], Agarwal et al. [1–3], Cecchi et al. [6], Parhi and Pardi [15], and the present authors [4, 8]. The suitable illustrative examples are also provided.

References

[1] R. P. Agarwal, S.-L. Shieh, and C.-C. Yeh, "Oscillation criteria for second-order retarded differential equations," *Mathematical and Computer Modelling*, vol. 26, no. 4, pp. 1–11, 1997.

[2] R. P. Agarwal, S. R. Grace, and D. O'Regan, "On the oscillation of certain functional differential equations via comparison methods," *Journal of Mathematical Analysis and Applications*, vol. 286, no. 2, pp. 577–600, 2003.

[3] R. P. Agarwal, S. R. Grace, and T. Smith, "Oscillation of certain third order functional differential equations," *Advances in Mathematical Sciences and Applications*, vol. 16, no. 1, pp. 69–94, 2006.

[4] B. Baculíková and J. Džurina, "Oscillation of third-order neutral differential equations," *Mathematical and Computer Modelling*, vol. 52, no. 1-2, pp. 215–226, 2010.

[5] B. Baculíková, R. P. Agarwal, T. Li, and J. Džurina, "Oscillation of third-order nonlinear functional differential equations with mixed arguments," to appear in *Acta Mathematica Hungarica*.

[6] M. Cecchi, Z. Došlá, and M. Marini, "On third order differential equations with property A and B," *Journal of Mathematical Analysis and Applications*, vol. 231, no. 2, pp. 509–525, 1999.

to Contents, page 1

[7] J. Džurina, "Asymptotic properties of third order delay differential equations," *Czechoslovak Mathematical Journal*, vol. 45(120), no. 3, pp. 443–448, 1995.

[8] J. Džurina, "Comparison theorems for functional-differential equations with advanced argument," *Unione Matematica Italiana. Bollettino*, vol. 7, no. 3, pp. 461–470, 1993.

[9] L. H. Erbe, Q. Kong, and B. G. Zhang, *Oscillation Theory for Functional-Differential Equations*, vol. 190 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 1994.

[10] S. R. Grace, R. P. Agarwal, R. Pavani, and E. Thandapani, "On the oscillation of certain third order nonlinear functional differential equations," *Applied Mathematics and Computation*, vol. 202, no. 1, pp. 102–112, 2008.

[11] I. Györi and G. Ladas, *Oscillation Theory of Delay Differential Equations*, Oxford Mathematical Monographs, The Clarendon Press, New York, NY, USA, 1991.

[12] T. S. Hassan, "Oscillation of third order nonlinear delay dynamic equations on time scales," *Mathematical and Computer Modelling*, vol. 49, no. 7-8, pp. 1573–1586, 2009.

[13] T. Kusano and M. Naito, "Comparison theorems for functional-differential equations with deviating arguments," *Journal of the Mathematical Society of Japan*, vol. 33, no. 3, pp. 509–532, 1981.

[14] G. S. Ladde, V. Lakshmikantham, and B. G. Zhang, *Oscillation Theory of Differential Equations with Deviating Arguments*, vol. 110 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 1987.

[15] N. Parhi and S. Pardi, "On oscillation and asymptotic property of a class of third order differential equations," *Czechoslovak Mathematical Journal*, vol. 49(124), no. 1, pp. 21–33, 1999.

[16] Ch. G. Philos, "On the existence of nonoscillatory solutions tending to zero at ∞ for differential equations with positive delays," *Archiv der Mathematik*, vol. 36, no. 2, pp. 168–178, 1981.

[17] A. Tiryaki and M. F. Aktaş, "Oscillation criteria of a certain class of third order nonlinear delay differential equations with damping," *Journal of Mathematical Analysis and Applications*, vol. 325, no. 1, pp. 54–68, 2007.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 919538, 35 pages
doi:10.1155/2011/919538

Research Article

The Lie Group in Infinite Dimension

V. Tryhuk, V. Chrastinová, and O. Dlouhý

Department of Mathematics, Faculty of Civil Engineering, Brno University of Technology, Veveří 331/95, 602 00 Brno, Czech Republic

Correspondence should be addressed to V. Tryhuk, tryhuk.v@fce.vutbr.cz

Received 6 December 2010; Accepted 12 January 2011

Academic Editor: Miroslava Růžičková

A Lie group acting on finite-dimensional space is generated by its infinitesimal transformations and conversely, any Lie algebra of vector fields in finite dimension generates a Lie group (the first fundamental theorem). This classical result is adjusted for the infinite-dimensional case. We prove that the (local, C^∞ smooth) action of a Lie group on infinite-dimensional space (a manifold modelled on $\mathbb{R}^\infty$) may be regarded as a limit of finite-dimensional approximations and the corresponding Lie algebra of vector fields may be characterized by certain finiteness requirements. The result is applied to the theory of generalized (or higher-order) infinitesimal symmetries of differential equations.

1. Preface

In the symmetry theory of differential equations, the *generalized* (*or*: *higher-order, Lie-Bäcklund*) *infinitesimal symmetries*

$$Z = \sum z_i \frac{\partial}{\partial x_i} + \sum z_I^j \frac{\partial}{\partial w_I^j} \quad \left(i = 1, \ldots, n;\ j = 1, \ldots, m;\ I = i_1 \cdots i_n;\ i_1, \ldots, i_n = 1, \ldots, n\right), \tag{1.1}$$

where the coefficients

$$z_i = z_i\left(\ldots, x_{i'}, w_{I'}^{j'}, \ldots\right), \qquad z_I^j = z_I^j\left(\ldots, x_{i'}, w_{I'}^{j'}, \ldots\right) \tag{1.2}$$

are functions of independent variables x_i, dependent variables w^j and a finite number of jet variables $w_I^j = \partial^n w^j / \partial x_{i_1} \cdots \partial x_{i_n}$ belong to well-established concepts. However, in spite of

to Contents, page 1

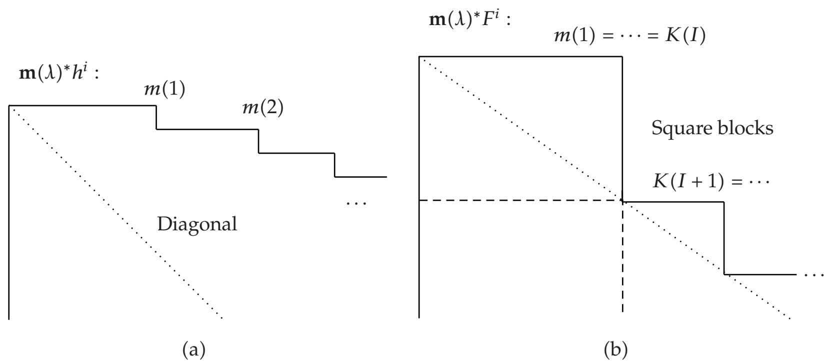


Figure 1

this matter of fact, they cause an unpleasant feeling. Indeed, such vector fields as a rule do not generate any one-parameter group of transformations

$$\overline{x}_i = G_i\left(\lambda; \ldots, x_{i'}, w_{I'}^{j'}, \ldots\right), \qquad \overline{w}_I^j = G_I^j\left(\lambda; \ldots, x_{i'}, w_{I'}^{j'}, \ldots\right) \tag{1.3}$$

in the underlying infinite-order jet space since the relevant Lie system

$$\frac{\partial G_i}{\partial \lambda} = z_i\left(\ldots, G_{i'}, G_{I'}^{j'}, \ldots\right), \quad \frac{\partial G_I^j}{\partial \lambda} = z_I^j\left(\ldots, G_{i'}, G_{I'}^{j'}, \ldots\right) \quad \left(G_i|_{\lambda=0} = x_i, \; G_I^j\Big|_{\lambda=0} = w_I^j\right) \tag{1.4}$$

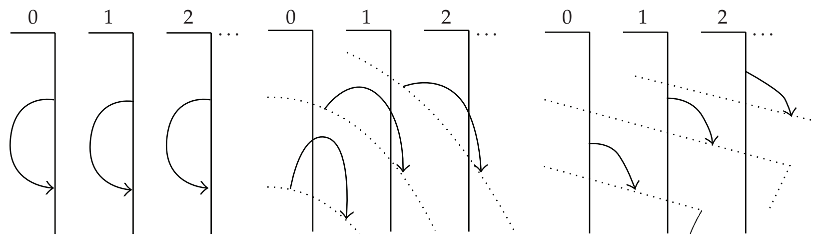
need not have any reasonable (locally unique) solution. Then Z is a mere formal concept [1–7] not related to any true transformations and the term "infinitesimal symmetry Z" is misleading, no Z-symmetries of differential equations in reality appear.

In order to clarify the situation, we consider one-parameter groups of local transformations in $\mathbb{R}^\infty$. We will see that they admit "finite-dimensional approximations" and as a byproduct, the relevant infinitesimal transformations may be exactly characterized by certain "finiteness requirements" of purely algebraical nature. With a little effort, the multidimensional groups can be easily involved, too. This result was briefly discussed in [8, page 243] and systematically mentioned at several places in monograph [9], but our aim is to make some details more explicit in order to prepare the necessary tools for systematic investigation of *groups of generalized symmetries*. We intend to continue our previous articles [10–13] where the algorithm for determination of all *individual generalized symmetries* was already proposed.

For the convenience of reader, let us transparently describe the crucial approximation result. We consider transformations (2.1) of a local one-parameter group in the space $\mathbb{R}^\infty$ with coordinates $h^1, h^2, \ldots$. Equations (2.1) of transformations $\mathbf{m}(\lambda)$ can be schematically represented by Figure 1(a).

We prove that in appropriate new coordinate system $F^1, F^2, \ldots$ on $\mathbb{R}^\infty$, the same transformations $\mathbf{m}(\lambda)$ become block triangular as in Figure 1(b). It follows that a certain hierarchy of finite-dimensional subspaces of $\mathbb{R}^\infty$ is preserved which provides the "approximation" of $\mathbf{m}(\lambda)$. The infinitesimal transformation $Z = \mathrm{d}\mathbf{m}(\lambda)/\mathrm{d}\lambda|_{\lambda=0}$ clearly preserves the same hierarchy which provides certain algebraical "finiteness" of Z.

to Contents, page 1



(a) (b) (c)

Figure 2

If the primary space $\mathbb{R}^\infty$ is moreover equipped with an appropriate structure, for example, the contact forms, it turns into the jet space and the results concerning the transformation groups on $\mathbb{R}^\infty$ become the theory of higher-order symmetries of differential equations. Unlike the common point symmetries which occupy a number of voluminous monographs (see, e.g., [14, 15] and extensive references therein) this higher-order theory was not systematically investigated yet. We can mention only the isolated article [16] which involves a direct proof of the "finiteness requirements" for one-parameter groups (namely, the result (ι) of Lemma 5.4 below) with two particular examples and monograph [7] involving a theory of generalized infinitesimal symmetries in the formal sense.

Let us finally mention the intentions of this paper. In the *classical theory of point or Lie's contact-symmetries* of differential equations, the order of derivatives is preserved (Figure 2(a)). Then the common Lie's and Cartan's methods acting in finite dimensional spaces *given ahead of calculations* can be applied. On the other extremity, the *generalized symmetries* need not preserve the order (Figure 2(c)) and even any finite-dimensional space and then the common classical methods fail. For the favourable intermediate case of *groups of generalized symmetries*, the invariant finite-dimensional subspaces exist, however, they are *not known in advance* (Figure 2(b)). We believe that the classical methods can be appropriately adapted for the latter case, and this paper should be regarded as a modest preparation for this task.

2. Fundamental Approximation Results

Our reasonings will be carried out in the space $\mathbb{R}^\infty$ with coordinates $h^1, h^2, \ldots$ [9] and we introduce the structural family $\mathcal{F}$ of all real-valued, locally defined and C^∞-smooth functions $f = f(h^1, \ldots, h^{m(f)})$ depending on a finite number of coordinates. In future, such functions will contain certain C^∞-smooth real parameters, too.

We are interested in (local) groups of transformations $\mathbf{m}(\lambda)$ in $\mathbb{R}^\infty$ defined by formulae

$$\mathbf{m}(\lambda)^* h^i = H^i\left(\lambda; h^1, \ldots, h^{m(i)}\right), \quad -\varepsilon^i < \lambda < \varepsilon^i,\ \varepsilon^i > 0\ (i = 1, 2, \ldots), \tag{2.1}$$

where $H^i \in \mathcal{F}$ if the parameter λ is kept fixed. We suppose

$$\mathbf{m}(0) = \text{id.}, \qquad \mathbf{m}(\lambda + \mu) = \mathbf{m}(\lambda)\mathbf{m}(\mu) \tag{2.2}$$

to Contents, page 1

whenever it makes a sense. An open and common definition domain for all functions H^i is tacitly supposed. (In more generality, a common definition domain for *every finite number* of functions H^i is quite enough and the germ and sheaf terminology would be more adequate for our reasonings, alas, it looks rather clumsy.)

Definition 2.1. For every $I = 1, 2, \ldots$ and $0 < \varepsilon < \min\{\varepsilon^1, \ldots, \varepsilon^I\}$, let $\mathcal{F}(I, \varepsilon) \subset \mathcal{F}$ be the subset of all composed functions

$$F = F\left(\ldots, \mathbf{m}(\lambda_j)^* h^i, \ldots\right) = F\left(\ldots, H^i\left(\lambda_j; h^1, \ldots, h^{m(i)}\right), \ldots\right), \tag{2.3}$$

where $i = 1, \ldots, I$; $-\varepsilon < \lambda_j < \varepsilon$; $j = 1, \ldots, J = J(I) = \max\{m(1), \ldots, m(I)\}$ and F is arbitrary C^∞-smooth function (of IJ variables). In functions $F \in \mathcal{F}(I, \varepsilon)$, variables $\lambda_1, \ldots, \lambda_J$ are regarded as mere parameters.

Functions (2.3) will be considered on open subsets of $\mathbb{R}^\infty$ where the rank of the Jacobi $(IJ \times J)$-matrix

$$\left(\frac{\partial}{\partial h^{j'}} H^i\left(\lambda_j; h^1, \ldots, h^{m(i)}\right)\right) \quad (i = 1, \ldots, I;\ j, j' = 1, \ldots, J) \tag{2.4}$$

of functions $H^i(\lambda_j; h^1, \ldots, h^{m(i)})$ locally attains the maximum (for appropriate choice of parameters). This rank and therefore the subset $\mathcal{F}(I, \varepsilon) \subset \mathcal{F}$ does not depend on ε as soon as $\varepsilon = \varepsilon(I)$ is close enough to zero. *This is supposed from now on and we may abbreviate* $\mathcal{F}(I) = \mathcal{F}(I, \varepsilon)$.

We deal with highly nonlinear topics. Then the definition domains cannot be kept fixed in advance. Our results will be true *locally*, near *generic points*, on certain *open everywhere dense subsets* of the underlying space $\mathbb{R}^\infty$. With a little effort, the subsets can be exactly characterized, for example, by locally constant rank of matrices, functional independence, existence of implicit function, and so like. We follow the common practice and as a rule omit such routine details from now on.

Lemma 2.2 (approximation lemma). *The following inclusion is true:*

$$\mathbf{m}(\lambda)^* \mathcal{F}(I) \subset \mathcal{F}(I). \tag{2.5}$$

Proof. Clearly

$$\mathbf{m}(\lambda)^* H^i(\lambda_j; \ldots) = \mathbf{m}(\lambda)^* \mathbf{m}(\lambda_j)^* h^i = \mathbf{m}(\lambda + \lambda_j)^* h^i = H^i(\lambda + \lambda_j; \ldots) \tag{2.6}$$

and therefore

$$\mathbf{m}(\lambda)^* F = F\left(\ldots, H^i\left(\lambda + \lambda_j; h^1, \ldots, h^{m(i)}\right), \ldots\right) \in \mathcal{F}(I). \tag{2.7}$$

□

to Contents, page 1

Denoting by $K(I)$ the rank of matrix (2.4), there exist *basical functions*

$$F^k = F^k\left(\dots, H^i\left(\lambda_j; h^1, \dots, h^{m(i)}\right), \dots\right) \in \mathcal{F}(I) \quad (k = 1, \dots, K(I)) \tag{2.8}$$

such that $\operatorname{rank}(\partial F^k / \partial h^{j'}) = K(I)$. Then a function $f \in \mathcal{F}$ lies in $\mathcal{F}(I)$ if and only if $f = \overline{f}(F^1, \dots, F^{K(I)})$ is a composed function. In more detail

$$F = \overline{F}\left(\lambda_1, \dots, \lambda_J; F^1, \dots, F^{K(I)}\right) \in \mathcal{F}(I) \tag{2.9}$$

is such a composed function if we choose $f = F$ given by (2.3). Parameters $\lambda_1, \dots, \lambda_J$ occurring in (2.3) are taken into account here. It follows that

$$\frac{\partial F}{\partial \lambda_j} = \frac{\partial \overline{F}}{\partial \lambda_j}\left(\lambda_1, \dots, \lambda_J; F^1, \dots, F^{K(I)}\right) \in \mathcal{F}(I) \quad (j = 1, \dots, J) \tag{2.10}$$

and analogously for the higher derivatives.

In particular, we also have

$$H^i\left(\lambda; h^1, \dots, h^{m(i)}\right) = \overline{H}^i\left(\lambda; F^1, \dots, F^{K(I)}\right) \in \mathcal{F}(I) \quad (i = 1, \dots, I) \tag{2.11}$$

for the choice $F = H^i(\lambda; \dots)$ in (2.9) whence

$$\frac{\partial^r H^i}{\partial \lambda^r} = \frac{\partial^r \overline{H}^i}{\partial \lambda^r}\left(\lambda; F^1, \dots, F^{K(I)}\right) \in \mathcal{F}(I) \quad (i = 1, \dots, I;\ r = 0, 1, \dots). \tag{2.12}$$

The basical functions can be taken from the family of functions $H^i(\lambda; \dots)$ $(i = 1, \dots, I)$ for appropriate choice of *various* values of λ. Functions (2.12) are enough as well even for a *fixed* value λ, for example, for $\lambda = 0$, see Theorem 3.2 below.

Lemma 2.3. *For any basical function, one has*

$$\mathbf{m}(\lambda)^* F^k = \overline{F}^k\left(\lambda; F^1, \dots, F^{K(I)}\right) \quad (k = 1, \dots, K(I)). \tag{2.13}$$

Proof. $F^k \in \mathcal{F}(I)$ implies $\mathbf{m}(\lambda)^* F^k \in \mathcal{F}(I)$ and (2.9) may be applied with the choice $F = \mathbf{m}(\lambda)^* F^k$ and $\lambda_1 = \cdots = \lambda_J = \lambda$. ☐

Summary 1. Coordinates $h^i = H^i(0; \dots)$ $(i = 1, \dots, I)$ were included into the subfamily $\mathcal{F}(I) \subset \mathcal{F}$ which is transformed into itself by virtue of (2.13). So we have a one-parameter group acting on $\mathcal{F}(I)$. One can even choose $F^1 = h^1, \dots, F^I = h^I$ here and then, if I is large enough, formulae (2.13) provide a "finite-dimensional approximation" of the primary mapping $\mathbf{m}(\lambda)$. The block-triangular structure of the infinite matrix of transformations $\mathbf{m}(\lambda)$ mentioned in Section 1 appears if $I \to \infty$ and the system of functions $F^1, F^2, \dots$ is succesively completed.

to Contents, page 1

3. The Infinitesimal Approach

We introduce the vector field

$$Z = \sum z^i \frac{\partial}{\partial h^i} = \frac{\mathrm{d}\mathbf{m}(\lambda)}{\mathrm{d}\lambda}\bigg|_{\lambda=0} \quad \left(z^i = \frac{\partial H^i}{\partial \lambda}\left(0; h^1, \ldots, h^{m(i)}\right);\ i = 1, 2, \ldots\right), \tag{3.1}$$

the *infinitesimal transformation* ($\mathcal{IT}$) of group $\mathbf{m}(\lambda)$. Let us recall the celebrated Lie system

$$\begin{aligned} \frac{\partial}{\partial \lambda}\mathbf{m}(\lambda)^* h^i &= \frac{\partial H^i}{\partial \lambda}(\lambda; \ldots) = \frac{\partial H^i}{\partial \mu}(\lambda + \mu; \ldots)\bigg|_{\mu=0} \\ &= \frac{\partial}{\partial \mu}\mathbf{m}(\lambda + \mu)^* h^i\bigg|_{\mu=0} = \mathbf{m}(\lambda)^* \frac{\partial}{\partial \mu}\mathbf{m}(\mu)^* h^i\bigg|_{\mu=0} = \mathbf{m}(\lambda)^* Z h^i = \mathbf{m}(\lambda)^* z^i. \end{aligned} \tag{3.2}$$

In more explicit (and classical) transcription

$$\frac{\partial H^i}{\partial \lambda}\left(\lambda; h^1, \ldots, h^{m(i)}\right) = z^i\left(H^1\left(\lambda; h^1, \ldots, h^{m(1)}\right), \ldots, H^{m(i)}\left(\lambda; h^1, \ldots, h^{m(m(i))}\right)\right). \tag{3.3}$$

One can also check the general identity

$$\frac{\partial^r}{\partial \lambda^r}\mathbf{m}(\lambda)^* f = \mathbf{m}(\lambda)^* Z^r f \quad (f \in \mathcal{F};\ r = 0, 1, \ldots) \tag{3.4}$$

by a mere routine induction on r.

Lemma 3.1 (finiteness lemma). *For all $r \in \mathbb{N}$, $Z^r \mathcal{F}(I) \subset \mathcal{F}(I)$.*

Proof. Clearly

$$ZF = \mathbf{m}(\lambda)^* ZF|_{\lambda=0} = \frac{\partial}{\partial \lambda}\mathbf{m}(\lambda)^* F\bigg|_{\lambda=0} \in \mathcal{F}(I) \tag{3.5}$$

for any function (2.3) by virtue of (2.10): induction on r. □

Theorem 3.2 (finiteness theorem). *Every function $F \in \mathcal{F}(I)$ admits (locally, near generic points) the representation*

$$F = \tilde{F}\left(\ldots, \frac{\partial^r H^i}{\partial \lambda^r}\left(0; h^1, \ldots, h^{m(i)}\right), \ldots\right) \tag{3.6}$$

in terms of a composed function where $i = 1, \ldots, I$ and $\tilde{F}$ is a $\mathbb{C}^\infty$-smooth function of a finite number of variables.

to Contents, page 1

Proof. Let us temporarily denote

$$H_r^i = \frac{\partial^r H^i}{\partial \lambda^r}(\lambda; \ldots) = \frac{\partial^r}{\partial \lambda^r}\mathbf{m}(\lambda)^* h^i, \qquad h_r^i = H_r^i(0; \ldots) = Z^r h^i, \tag{3.7}$$

where the second equality follows from (3.4) with $f = h^i$, $\lambda = 0$. Then

$$H_r^i = \mathbf{m}(\lambda)^* h_r^i = \mathbf{m}(\lambda)^* Z^r h^i \tag{3.8}$$

by virtue of (3.4) with general λ.

If $j = j(i)$ is large enough, there does exist an identity $h_{j+1}^i = G^i(h_0^i, \ldots, h_j^i)$. Therefore

$$\frac{\partial^{j+1} H^i}{\partial \lambda^{j+1}} = H_{j+1}^i = G^i\left(H_0^i, \ldots, H_j^i\right) = G^i\left(H^i, \ldots, \frac{\partial^j H^i}{\partial \lambda^j}\right) \tag{3.9}$$

by applying $\mathbf{m}(\lambda)^*$. This may be regarded as ordinary differential equation with initial values

$$H^i\Big|_{\lambda=0} = h_0^i, \ldots, \frac{\partial^j H^i}{\partial \lambda^j}\Big|_{\lambda=0} = h_j^i. \tag{3.10}$$

The solution $H^i = \widetilde{H}^i(\lambda; h_0^i, \ldots, h_j^i)$ expressed in terms of initial values reads

$$H^i\left(\lambda; h^1, \ldots, h^{m(i)}\right) = \widetilde{H}^i\left(\lambda; H^i\left(0; h^1, \ldots, h^{m(i)}\right), \ldots, \frac{\partial^j H^i}{\partial \lambda^j}\left(0; h^1, \ldots, h^{m(i)}\right)\right) \tag{3.11}$$

in full detail. If λ is kept fixed, this is exactly the identity (3.6) for the particular case $F = H^i(\lambda; h^1, \ldots, h^{m(i)})$. The general case follows by a routine. □

Definition 3.3. Let $\mathbb{G}$ be the set of (local) vector fields

$$Z = \sum z^i \frac{\partial}{\partial h^i} \quad \left(z^i \in \mathcal{F}, \text{ infinite sum}\right) \tag{3.12}$$

such that every family of functions $\{Z^r h^i\}_{r\in\mathbb{N}}$ (i fixed but arbitrary) can be expressed in terms of a finite number of coordinates.

Remark 3.4. Neither $\mathbb{G} + \mathbb{G} \subset \mathbb{G}$ nor $[\mathbb{G}, \mathbb{G}] \subset \mathbb{G}$ as follows from simple examples. However, $\mathbb{G}$ is a *conical set* (over $\mathcal{F}$): if $Z \in \mathbb{G}$ then $fZ \in \mathbb{G}$ for any $f \in \mathcal{F}$. Easy direct proof may be omitted here.

Summary 2. If Z is $\mathcal{IT}$ of a group then all functions $Z^r h^i$ ($i = 1, \ldots, I$; $r = 0, 1, \ldots$) are included into family $\mathcal{F}(I)$ hence $Z \in \mathbb{G}$. The converse is clearly also true: every vector field $Z \in \mathbb{G}$ generates a local Lie group since the Lie system (3.3) admits finite-dimensional approximations in spaces $\mathcal{F}(I)$.

to Contents, page 1

Let us finally reformulate the last sentence in terms of basical functions.

Theorem 3.5 (approximation theorem). *Let $Z \in \mathbb{G}$ be a vector field locally defined on $\mathbb{R}^\infty$ and $F^1, \ldots, F^{K(I)} \in \mathcal{F}$ be a maximal functionally independent subset of the family of all functions*

$$Z^r h^i \quad (i = 1, \ldots, I;\ r = 0, 1, \ldots). \tag{3.13}$$

Denoting $ZF^k = \overline{F}^k(F^1, \ldots, F^{K(I)})$, then the system

$$\frac{\partial}{\partial \lambda}\mathbf{m}(\lambda)^* F^k = \mathbf{m}(\lambda)^* Z F^k = \overline{F}^k\left(\mathbf{m}(\lambda)^* F^1, \ldots, \mathbf{m}(\lambda)^* F^{K(I)}\right) \quad (k = 1, \ldots, K(I)) \tag{3.14}$$

may be regarded as a "finite-dimensional approximation" to the Lie system (3.3) *of the one-parameter local group* $\mathbf{m}(\lambda)$ *generated by Z.*

In particular, assuming $F^1 = h^1, \ldots, F^I = h^I$, then the the initial portion

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbf{m}(\lambda)^* F^i = \frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbf{m}(\lambda)^* h^i = \frac{\mathrm{d}}{\mathrm{d}\lambda} H^i = z^i\left(H^1, \ldots, H^{m(i)}\right) \quad (i = 1, \ldots, I) \tag{3.15}$$

of the above system transparently demonstrates the approximation property.

4. On the Multiparameter Case

The following result does not bring much novelty and we omit the proof.

Theorem 4.1. *Let $Z_1, \ldots, Z_d$ be commuting local vector fields in the space $\mathbb{R}^\infty$. Then $Z_1, \ldots, Z_d \in \mathbb{G}$ if and only if the vector fields $Z = a_1 Z_1 + \cdots + a_d Z_d$ $(a_1, \ldots, a_d \in \mathbb{R})$ locally generate an abelian Lie group.*

In full non-Abelian generality, let us consider a (local) multiparameter group formally given by the same equations (2.1) as above where $\lambda = (\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^d$ are parameters close to the zero point $0 = (0, \ldots, 0) \in \mathbb{R}^d$. The rule (2.2) is generalized as

$$\mathbf{m}(0) = \mathrm{id.}, \qquad \mathbf{m}(\varphi(\lambda, \mu)) = \mathbf{m}(\lambda)\mathbf{m}(\mu), \tag{4.1}$$

where $\lambda = (\lambda_1, \ldots, \lambda_d)$, $\mu = (\mu_1, \ldots, \mu_d)$ and $\varphi = (\varphi_1, \ldots, \varphi_d)$ determine the composition of parameters. Appropriately adapting the space $\mathcal{F}(I)$ and the concept of basical functions $F^1, \ldots, F^{K(I)}$, Lemma 2.2 holds true without any change.

Passing to the infinitesimal approach, we introduce vector fields $Z_1, \ldots, Z_d$ which are $\mathcal{IT}$ of the group. We recall (without proof) the Lie equations [17]

$$\frac{\partial}{\partial \lambda_j}\mathbf{m}(\lambda)^* f = \sum a_i^j(\lambda)\mathbf{m}(\lambda)^* Z_j f \quad (f \in \mathcal{F};\ j = 1, \ldots, d) \tag{4.2}$$

with the initial condition $\mathbf{m}(0) = \mathrm{id}$. Assuming $Z_1, \ldots, Z_d$ linearly independent over $\mathbb{R}$, coefficients $a_i^j(\lambda)$ may be arbitrarily chosen and the solution $\mathbf{m}(\lambda)$ always is a group

to Contents, page 1

transformation (the first fundamental theorem). If basical functions $F^1,\ldots,F^{K(I)}$ are inserted for f, we have a finite-dimensional approximation which is self-contained in the sense that

$$Z_j F^k = \tilde{F}^k_j\left(F^1,\ldots,F^{K(I)}\right) \quad (j=1,\ldots,d;\ k=1,\ldots,K(I)) \tag{4.3}$$

are composed functions in accordance with the definition of the basical functions.

Let us conversely consider a Lie algebra of local vector fields $Z = a_1Z_1+\cdots+a_dZ_d$ $(a_i \in \mathbb{R})$ on the space $\mathbb{R}^\infty$. Let moreover $Z_1,\ldots,Z_d \in \mathbb{G}$ *uniformly* in the sense that there is a *universal space* $\mathcal{F}(I)$ with $\mathcal{L}_{Z_i}\mathcal{F}(I) \subset \mathcal{F}(I)$ for all $i=1,\ldots,d$. Then the Lie equations may be applied and we obtain reasonable finite-dimensional approximations.

Summary 3. Theorem 4.1 holds true even in the non-Abelian and multidimensional case if the inclusions $Z_1,\ldots,Z_d \in \mathbb{G}$ are uniformly satisfied.

As yet we have closely simulated the primary one-parameter approach, however, the results are a little misleading: the uniformity requirement in Summary 3 may be completely omitted. This follows from the following result [9, page 30] needless here and therefore stated without proof.

Theorem 4.2. *Let $\mathcal{K}$ be a finite-dimensional submodule of the module of vector fields on $\mathbb{R}^\infty$ such that $[\mathcal{K},\mathcal{K}] \subset \mathcal{K}$. Then $\mathcal{K} \subset \mathbb{G}$ if and only if there exist generators (over $\mathcal{F}$) of submodule $\mathcal{K}$ that are lying in $\mathbb{G}$.*

5. Symmetries of the Infinite-Order Jet Space

The previous results can be applied to the groups of generalized symmetries of *partial differential equations*. Alas, some additional technical tools cannot be easily explained at this place, see the concluding Section 11 below. So we restrict ourselves to the *trivial differential equations*, that is, to the groups of generalized symmetries in the total *infinite-order jet space* which do not require any additional preparations.

Let $\mathbf{M}(m,n)$ be the jet space of n-dimensional submanifolds in $\mathbb{R}^{m+n}$ [9–13]. We recall the familiar (local) jet coordinates

$$x_i, w^j_I \quad (I = i_1\ldots i_r;\ i,i_1,\ldots,i_r = 1,\ldots,n;\ r=0,1,\ldots;\ j=1,\ldots,m). \tag{5.1}$$

Functions $f = f(\ldots,x_i,w^j_I,\ldots)$ on $\mathbf{M}(m,n)$ are C^∞-smooth and depend on a finite number of coordinates. The jet coordinates serve as a mere technical tool. The true jet structure is given just by the *module* $\Omega(m,n)$ *of contact forms*

$$\omega = \sum a^j_I \omega^j_I \quad \left(\text{finite sum}, \omega^j_I = \mathrm{d}w^j_I - \sum w^j_{Ii}\mathrm{d}x_i\right) \tag{5.2}$$

or, equivalently, by the *"orthogonal" module* $\mathcal{H}(m,n) = \Omega^\perp(m,n)$ *of formal derivatives*

$$D = \sum a_i D_i \left(D_i = \frac{\partial}{\partial x_i} + \sum w^j_{Ii}\frac{\partial}{\partial w^j_I};\ i=1,\ldots,n; D\rfloor\omega^j_I = \omega^j_I(D) = 0\right). \tag{5.3}$$

Let us state useful formulae

$$\mathrm{d}f = \sum D_i f \,\mathrm{d}x_i + \sum \frac{\partial f}{\partial w_I^j} \omega_I^j, \qquad D_i \rfloor \mathrm{d}\omega_I^j = \omega_{Ii}^j, \qquad \mathcal{L}_{D_i} \omega_I^j = \omega_{Ii}^j, \tag{5.4}$$

where $\mathcal{L}_{D_i} = D_i \rfloor \mathrm{d} + \mathrm{d}D_i \rfloor$ denotes the Lie derivative.

We are interested in (local) one-parameter groups of transformations $\mathbf{m}(\lambda)$ given by certain formulae

$$\mathbf{m}(\lambda)^* x_i = G_i\left(\lambda; \ldots, x_{i'}, w_{I'}^{j'}, \ldots\right), \qquad \mathbf{m}(\lambda)^* w_I^j = G_I^j\left(\lambda; \ldots, x_{i'}, w_{I'}^{j'}, \ldots\right) \tag{5.5}$$

and in vector fields

$$Z = \sum z_i\left(\ldots, x_{i'}, w_{I'}^{j'}, \ldots\right) \frac{\partial}{\partial x_i} + \sum z_I^j\left(\ldots, x_{i'}, w_{I'}^{j'}, \ldots\right) \frac{\partial}{\partial w_I^j} \tag{5.6}$$

locally defined on the jet space $\mathbf{M}(m, n)$; see also (1.1) and (1.2).

Definition 5.1. We speak of a *group of morphisms* (5.5)of the jet structure if the inclusion $\mathbf{m}(\lambda)^* \Omega(m, n) \subset \Omega(m, n)$ holds true. We speak of a *(universal) variation* (5.6) of the jet structure if$\mathcal{L}_Z \Omega(m, n) \subset \Omega(m, n)$. If a variation (5.6) moreover generates a group, speaks of a *(generalized or higher-order) infinitesimal symmetry* of the jet structure.

So we intentionally distinguish between true infinitesimal transformations generating a group and the formal concepts; this point of view and the terminology are not commonly used in the current literature.

Remark 5.2. A few notes concerning this unorthodox terminology are useful here. In actual literature, the vector fields (5.6) are as a rule decomposed into the "trivial summand D" and the so-called "evolutionary form V" of the vector field Z, explicitly

$$Z = D + V \quad \left(D = \sum z_i D_i \in \mathcal{H}(m, n), V = \sum Q_I^j \frac{\partial}{\partial w_I^j},\ Q_I^j = z_I^j - \sum w_{Ii}^j z_i \right). \tag{5.7}$$

The summand D is usually neglected in a certain sense [3–7] and the "essential" summand V is identified with the evolutional system

$$\frac{\partial w_I^j}{\partial \lambda} = Q_I^j\left(\ldots, x_{i'}, w_{I'}^{j'}, \ldots\right) \quad \left(w_I^j = \frac{\partial^n w^j}{\partial x_{i_1} \cdots \partial x_{i_n}} (\lambda, x_1, \ldots, x_n) \right) \tag{5.8}$$

of partial differential equations (the finite subsystem with $I = \phi$ empty is enough here since the remaining part is a mere prolongation). This evolutional system is regarded as a "virtual flow" on the "space of solutions" $w^j = w^j(x_1, \ldots, x_n)$, see [7, especially page 11]. In more generality, some differential constraints may be adjoint. However, in accordance with the

to Contents, page 1

ancient classical tradition, functions $\delta w^j = \partial w^j/\partial\lambda$ are just the *variations*. (There is only one novelty: in classical theory, δw^j are introduced only *along a given solution* while the vector fields Z are "universally" defined on the space.) In this "evolutionary approach", the properties of the primary vector field Z are utterly destroyed. It seems that the true sense of this approach lies in the applications to the topical soliton theory. However, then the evolutional system is always completed with boundary conditions and embedded into some normed functional spaces in order to ensure the existence of global "true flows". This is already quite a different story and we return to our topic.

In more explicit terms, morphisms (5.5) are characterized by the (implicit) recurrence

$$\sum G^j_{Ii} D_{i'} G_i = D_{i'} G^j_I \quad (i' = 1, \ldots, n), \tag{5.9}$$

where $\det(D_{i'}G_i) \neq 0$ is supposed and vector field (5.6) is a variation if and only if

$$z^j_{Ii} = D_i z^j_I - \sum w^j_{Ii'} D_i z_{i'}. \tag{5.10}$$

Recurrence (5.9) easily follows from the inclusion $\mathbf{m}(\lambda)^* \omega^j_I \in \Omega(m,n)$ and we omit the proof. Recurrence (5.10) follows from the identity

$$\begin{aligned}\mathcal{L}_Z \omega^j_I &= \mathcal{L}_Z \Big(\mathrm{d}w^j_I - \sum w^j_{Ii} \mathrm{d}x_i \Big) = \mathrm{d}z^j_I - \sum z^j_{Ii} \mathrm{d}x_i - \sum w^j_{Ii} \mathrm{d}z_i \\ &\cong \Big(\sum D_{i'} z^j_I - \sum z^j_{Ii'} - \sum w^j_{Ii} D_{i'} z_i \Big) \mathrm{d}x_{i'} \ (\mathrm{mod}\ \Omega(m,n))\end{aligned} \tag{5.11}$$

and the inclusion $\mathcal{L}_Z \omega^j_I \in \Omega(m,n)$. The obvious formula

$$\mathcal{L}_Z \omega^j_I = \sum \left(\frac{\partial z^j_I}{\partial w^{j'}_{I'}} - \sum w^j_{Ii} \frac{\partial z_i}{\partial w^{j'}_{I'}} \right) \omega^{j'}_{I'} \tag{5.12}$$

appearing on this occasion also is of a certain sense, see Theorem 5.5 and Section 10 below. It follows that the initial functions G_i, G^j, z_i, z^j (empty $I = \phi$) may be in principle arbitrarily prescribed in advance. This is the familiar *prolongation procedure* in the jet theory.

Remark 5.3. Recurrence (5.10) for the variation Z can be succintly expressed by $\omega^j_{Ii}(Z) = D_i \omega^j_I(Z)$. This remarkable formula admits far going generalizations, see concluding Examples 11.3 and 11.4 below.

Let us recall that a vector field (5.6) generates a group (5.5) if and only if $Z \in \mathbb{G}$ hence if and only if every family

$$\{Z^r x_i\}_{r\in\mathbb{N}}, \qquad \left\{Z^r w^j_I\right\}_{r\in\mathbb{N}} \tag{5.13}$$

to Contents, page 1

can be expressed in terms of a finite number of jet coordinates. We conclude with simple but practicable remark: due to jet structure, the *infinite number of conditions* (5.13) can be replaced by a *finite number of requirements* if Z is a variation.

Lemma 5.4. *Let* (5.6) *be a variation of the jet structure. Then the inclusion $Z \in \mathbb{G}$ is equivalent to any of the requirements*

(ι) *every family of functions*

$$\{Z^r x_i\}_{r\in\mathbb{N}}, \quad \left\{Z^r w^j\right\}_{r\in\mathbb{N}} \quad (i = 1,\dots,n;\ j = 1,\dots,m) \tag{5.14}$$

can be expressed in terms of a finite number of jet coordinates,

($\iota\iota$) *every family of differential forms*

$$\{\mathcal{L}_Z^r \mathrm{d}x_i\}_{r\in\mathbb{N}}, \quad \left\{\mathcal{L}_Z^r \mathrm{d}w^j\right\}_{r\in\mathbb{N}} \quad (i = 1,\dots,n;\ j = 1,\dots,m) \tag{5.15}$$

involves only a finite number of linearly independent terms,

($\iota\iota\iota$) *every family of differential forms*

$$\{\mathcal{L}_Z^r \mathrm{d}x_i\}_{r\in\mathbb{N}}, \quad \left\{\mathcal{L}_Z^r \mathrm{d}w_I^j\right\}_{r\in\mathbb{N}} \quad (i = 1,\dots,n;\ j = 1,\dots,m;\ \text{arbitrary } I) \tag{5.16}$$

involves only a finite number of linearly independent terms.

Proof. Inclusion $Z \in \mathbb{G}$ is defined by using the families (5.13) and this trivially implies (ι) where only the empty multi-indice $I = \phi$ is involved. Then (ι) implies ($\iota\iota$) by using the rule $\mathcal{L}_Z \mathrm{d}f = \mathrm{d}Zf$. Assuming ($\iota\iota$), we may employ the commutative rule

$$[D_i, Z] = D_i Z - Z D_i = \sum a_i^{i'} D_{i'} \quad \left(a_i^{i'} = D_i z_{i'}\right) \tag{5.17}$$

in order to verify identities of the kind

$$\mathcal{L}_Z \mathrm{d}w_i^j = \mathcal{L}_Z \mathrm{d}D_i w^j = \mathcal{L}_Z \mathcal{L}_{D_i} \mathrm{d}w^i = \mathcal{L}_{D_i} \mathcal{L}_Z \mathrm{d}w^i - \sum a_i^{i'} \mathcal{L}_{D_{i'}} w^j \tag{5.18}$$

and in full generality identities of the kind

$$\mathcal{L}_Z^k \mathrm{d}w_I^j = \sum a_{I,k}^{I'} \mathcal{L}_{D_{I'}} \mathcal{L}_Z^{k'} \mathrm{d}w^j \quad \left(\text{sum with } k' \le k,\ \left|I'\right| \le |I|\right) \tag{5.19}$$

with unimportant coefficients, therefore ($\iota\iota\iota$) follows. Finally ($\iota\iota\iota$) obviously implies the primary requirement on the families (5.13). □

This is not a whole story. The requirements can be expressed only in terms of the structural contact forms. With this final result, the algorithms [10–13] for determination of all *individual morphisms* can be closely simulated in order to obtain the algorithm for the determination of all *groups* $\mathbf{m}(\lambda)$ *of morphisms*, see Section 10 below.

to Contents, page 1

Theorem 5.5 (technical theorem). *Let* (5.6) *be a variation of the jet space. Then* $Z \in \mathbb{G}$ *if and only if every family*

$$\left\{ \mathcal{L}_Z^r \omega^j \right\}_{r \in \mathbb{N}} \quad (j = 1, \ldots, m) \tag{5.20}$$

involves only a finite number of linearly independent terms.

Some nontrivial preparation is needful for the proof. Let Θ be a finite-dimensional module of 1-forms (on the space $\mathbf{M}(m,n)$ but the underlying space is irrelevant here). Let us consider vector fields X such that $\mathcal{L}_{fX}\Theta \subset \Theta$ for all functions f. Let moreover Adj Θ be the module of all forms φ satisfying $\varphi(X) = 0$ for all such X. Then Adj Θ has a basis consisting of total differentials of certain functions $f^1, \ldots, f^K$ (the Frobenius theorem), and there is a basis of module Θ which can be expressed in terms of functions $f^1, \ldots, f^K$. Alternatively saying, (an appropriate basis of) the Pfaffian system $\vartheta = 0$ $(\vartheta \in \Theta)$ can be expressed only in terms of functions $f^1, \ldots, f^K$. This result frequently appears in Cartan's work, but we may refer only to [9, 18, 19] and to the appendix below for the proof.

Module Adj Θ is intrinsically related to Θ: if a mapping $\mathbf{m}$ preserves Θ then $\mathbf{m}$ preserves Adj Θ. In particular, assuming

$$\mathbf{m}(\lambda)^*\Theta \subset \Theta, \quad \text{then } \mathbf{m}(\lambda)^*\mathrm{Adj}\ \Theta \subset \mathrm{Adj}\ \Theta \tag{5.21}$$

is true for a group $\mathbf{m}(\lambda)$. In terms of $\mathcal{IT}$ of the group $\mathbf{m}(\lambda)$, we have equivalent assertion

$$\mathcal{L}_Z\Theta \subset \Theta \text{ implies } \mathcal{L}_Z\mathrm{Adj}\ \Theta \subset \mathrm{Adj}\ \Theta \tag{5.22}$$

and therefore $\mathcal{L}_Z^r\mathrm{Adj}\ \Theta \subset \mathrm{Adj}\ \Theta$ for all r. The preparation is done.

Proof. Let Θ be the module generated by all differential forms $\mathcal{L}_Z^r\omega^j$ $(j = 1, \ldots, m; r = 0, 1, \ldots)$. Assuming finite dimension of module Θ, we have module Adj Θ and clearly $\mathcal{L}_Z\Theta \subset \Theta$ whence $\mathcal{L}_Z^r\mathrm{Adj}\ \Theta \subset \mathrm{Adj}\ \Theta$ $(r = 0, 1, \ldots)$. However Adj Θ involves both the differentials $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ (see below) and the forms $\omega^1, \ldots, \omega^m$. Point (u) of previous Lemma 5.4 implies $Z \in \mathbb{G}$. The converse is trivial.

In order to finish the proof, let us on the contrary assume that Adj Θ *does not* contain all differentials $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$. Alternatively saying, the Pfaffian system $\vartheta = 0$ $(\vartheta \in \Theta)$ can be expressed in terms of certain functions $f^1, \ldots, f^K$ such that $\mathrm{d}f^1 = \cdots = \mathrm{d}f^K = 0$ does not imply $\mathrm{d}x_1 = \cdots = \mathrm{d}x_n = 0$. On the other hand, it follows clearly that maximal solutions of the Pfaffian system can be expressed only in terms of functions $f^1, \ldots, f^K$ and therefore we *do not need all* independent variables $x_1, \ldots, x_n$. This is however a contradiction: the Pfaffian system consists of contact forms and involves the equations $\omega^1 = \cdots = \omega^n = 0$. All independent variables are needful if we deal with the common classical solutions $w^j = w^j(x_1, \ldots, x_n)$. ☐

The result can be rephrased as follows.

Theorem 5.6. *Let* $\Omega_0 \subset \Omega(m,n)$ *be the submodule of all zeroth-order contact forms* $\omega = \sum a^j\omega^j$ *and* Z *be a variation of the jet structure. Then* $Z \in \mathbb{G}$ *if and only if* $\dim \oplus \mathcal{L}_Z^r\Omega_0 < \infty$.

to Contents, page 1

6. On the Multiparameter Case

Let us temporarily denote by $\mathbb{V}$ the family of all infinitesimal variations (5.6) of the jet structure. Then $\mathbb{V} + \mathbb{V} \subset \mathbb{V}$, $c\mathbb{V} \subset \mathbb{V}$ $(c \in \mathbb{R})$, $[\mathbb{V}, \mathbb{V}] \subset \mathbb{V}$, and it follows that $\mathbb{V}$ is an infinite-dimensional Lie algebra (coefficients in $\mathbb{R}$). On the other hand, if $Z \in \mathbb{V}$ and $fZ \in \mathbb{V}$ for certain $f \in \mathcal{F}$ then $f \in \mathbb{R}$ is a constant. (Briefly saying: *the conical variations of the total jet space do not exist*. We omit easy direct proof.) It follows that only the common Lie algebras over $\mathbb{R}$ are engaged if we deal with morphisms of the jet spaces $\mathbf{M}(m, n)$.

Theorem 6.1. *Let $\mathcal{G} \subset \mathbb{V}$ be a finite-dimensional Lie subalgebra. Then $\mathcal{G} \subset \mathbb{G}$ if and only if there exists a basis of $\mathcal{G}$ that is lying in $\mathbb{G}$.*

The proof is elementary and may be omitted. Briefly saying, Theorem 4.2 (coefficients in $\mathcal{F}$) turns into quite other and much easier Theorem 6.1 (coefficients in $\mathbb{R}$).

7. The Order-Preserving Groups in Jet Space

Passing to particular examples from now on, we will briefly comment some well-known classical results for the sake of completeness.

Let $\Omega_l \subset \Omega(m, n)$ be the submodule of all contact forms $\omega = \sum a_I^j \omega_I^j$ (sum with $|I| \le l$) of the order l at most. A morphism (5.5) and the infinitesimal variation (5.6) are called *order preserving* if

$$\mathbf{m}(\lambda)^* \Omega_l \subset \Omega_l, \qquad \mathcal{L}_Z \Omega_l \subset \Omega_l, \tag{7.1}$$

respectively, for a certain $l = 0, 1, \ldots$ (equivalently: for all $l \in \mathbb{N}$, see Lemmas 9.1 and 9.2 below). Due to the fundamental Lie-Bäcklund theorem [1, 3, 6, 10–13], this is possible only in the *pointwise case* or in the *Lie's contact transformation case*. In quite explicit terms: assuming (7.1) then either functions G_i, G^j, z_i, z^j (empty $I = \phi$) in formulae (5.5) and (5.6) are functions only of the zeroth-order jet variables $x_{i'}, w^{j'}$ or, in the second case, we have $m = 1$ and all functions $G_i, G^1, G_i^1, z_i, z^1, z_i^1$ contain only the zeroth- and first-order variables $x_{i'}, w^1, w_{i'}^1$.

A somewhat paradoxically, short proofs of this fundamental result are not easily available in current literature. We recall a tricky approach here already applied in [10–13], to the case of the order-preserving morphisms. The approach is a little formally improved and appropriately adapted to the infinitesimal case.

Theorem 7.1 (infinitesimal Lie-Bäcklund). *Let a variation Z preserve a submodule $\Omega_l \subset \Omega(m, n)$ of contact forms of the order l at most for a certain $l \in \mathbb{N}$. Then $Z \in \mathbb{G}$ and either Z is an infinitesimal point transformation or $m = 1$ and Z is the infinitesimal Lie's contact transformation.*

Proof. We suppose $\mathcal{L}_Z \Omega_l \subset \Omega_l$. Then $\mathcal{L}_Z^r \Omega_0 \subset \mathcal{L}_Z^r \Omega_l \subset \Omega_l$ therefore $Z \in \mathbb{G}$ by virtue of Theorem 5.5. Moreover $\mathcal{L}_Z \Omega_{l-1} \subset \Omega_{l-1}, \ldots, \mathcal{L}_Z \Omega_0 \subset \Omega_0$ by virtue of Lemma 9.2 below. So we have

$$\mathcal{L}_Z \omega^j = \sum a^{jj'} \omega^{j'} \quad (j, j' = 1, \ldots, m). \tag{7.2}$$

to Contents, page 1

Assuming $m = 1$, then (7.2) turns into the classical definition of Lie's infinitesimal contact transformation. Assume $m \geq 2$. In order to finish the proof we refer to the following result which implies that Z is indeed an infinitesimal point transformation. □

Lemma 7.2. *Let Z be a vector field on the jet space $\mathbf{M}(m,n)$ satisfying* (7.2) *and $m \geq 2$. Then*

$$Zx_i = z_i\left(\dots, x_{i'}, w^{j'}, \dots\right), \quad Zw^j = z^j\left(\dots, x_{i'}, w^{j'}, \dots\right) \quad (i = 1, \dots, n;\ j = 1, \dots, m) \tag{7.3}$$

are functions only of the point variables.

Proof. Let us introduce module Θ of $(m+2n)$-forms generated by all forms of the kind

$$\begin{aligned} &\omega^1 \wedge \cdots \wedge \omega^m \wedge \left(\mathrm{d}\omega^{j_1}\right)^{n_1} \wedge \left(\mathrm{d}\omega^{j_k}\right)^{n_k} \\ &\quad = \mathrm{d}w^1 \wedge \cdots \mathrm{d}w^m \wedge \mathrm{d}x_1 \wedge \cdots \mathrm{d}x_n \wedge \sum \pm\, \mathrm{d}w^{j'_1}_{i_1} \wedge \cdots \wedge \mathrm{d}w^{j'_n}_{i_n}, \end{aligned} \tag{7.4}$$

where $\sum n_k = n$. Clearly $\Theta = (\Omega_0)^m \wedge (\mathrm{d}\Omega_0)^n$. The inclusions

$$\mathcal{L}_Z\Omega_0 \subset \Omega_0, \qquad \mathcal{L}_Z\mathrm{d}\Omega_0 = \mathrm{d}\mathcal{L}_Z\Omega_0 + \Omega_0 \subset \mathrm{d}\Omega_0 + \Omega_0 \tag{7.5}$$

are true by virtue of (7.2) and imply $\mathcal{L}_Z\Theta \subset \Theta$.

Module Θ vanishes when restricted to certain hyperplanes, namely, just to the hyperplanes of the kind

$$\vartheta = \sum a_i \mathrm{d}x_i + \sum a^j \mathrm{d}w^j = 0 \tag{7.6}$$

(use $m \geq 2$ here). This is expressed by $\Theta \wedge \vartheta = 0$ and it follows that

$$0 = \mathcal{L}_Z(\Theta \wedge \vartheta) = \mathcal{L}_Z\Theta \wedge \vartheta + \Theta \wedge \mathcal{L}_Z\vartheta = \Theta \wedge \mathcal{L}_Z\vartheta. \tag{7.7}$$

Therefore $\mathcal{L}_Z\vartheta$ again is such a hyperplane: $\mathcal{L}_Z\vartheta \cong 0$ (mod all $\mathrm{d}x_i$ and $\mathrm{d}w^j$). On the other hand,

$$\mathcal{L}_Z\vartheta \cong \sum a_i \mathrm{d}z_i + \sum a^j \mathrm{d}z^j \quad \left(\text{mod all } \mathrm{d}x_i \text{ and } \mathrm{d}w^j\right) \tag{7.8}$$

and it follows that $\mathrm{d}z_i, \mathrm{d}z^j \cong 0$. □

There is a vast literature devoted to the pointwise transformations and symmetries so that any additional comments are needless. On the other hand, the contact transformations are more involved and less popular. They explicitly appear on rather peculiar and dissimilar occasions in actual literature [20, 21]. However, in reality the groups of Lie contact transformations are latently involved in the classical calculus of variations and provide the core of the Hilbert-Weierstrass extremality theory of variational integrals.

to Contents, page 1

8. Digression to the Calculus of Variations

We establish the following principle.

Theorem 8.1 (metatheorem). *The geometries of nondegenerate local one-parameter groups of Lie contact transformations* ($\mathcal{CT}$) *and of nondegenerate first-order one-dimensional variational integrals* ($\mathcal{VI}$) *are identical. In particular, the orbits of a given* $\mathcal{CT}$ *group are extremals of appropriate* $\mathcal{VI}$ *and conversely.*

Proof. The $\mathcal{CT}$ groups act in the jet space $\mathbf{M}(1,n)$ equipped with the contact module $\Omega(1,n)$. Then the abbreviations

$$w_I = w_I^1, \quad \omega_I = \omega_I^1 = \mathrm{d}w_I - \sum w_{Ii}\mathrm{d}x_i \quad Z = \sum z_i \frac{\partial}{\partial x_i} + \sum z_I^1 \frac{\partial}{\partial w_I} \tag{8.1}$$

are possible. Let us recall the classical approach [22, 23]. The Lie contact transformations defined by certain formulae

$$\mathbf{m}^* x_i = G_i(\cdot), \quad \mathbf{m}^* w = G^1(\cdot), \quad \mathbf{m}^* w_i = G_i^1(\cdot) \quad ((\cdot) = (x_1,\ldots,x_n,w,w_1,\ldots,w_n)) \tag{8.2}$$

preserve the Pfaffian equation $\omega = \mathrm{d}w - \sum w_i \mathrm{d}x_i = 0$ or (equivalently) the submodule $\Omega_0 \subset \Omega(1,n)$ of zeroth-order contact forms. Explicit formulae are available in literature. We are interested in one-parameter local $\mathcal{CT}$ groups of transformations $\mathbf{m}(\lambda)(-\varepsilon < \lambda < \varepsilon)$ which are "nondegenerate" in a sense stated below and then the explicit formulae are not available yet. On the other hand, our $\mathcal{VI}$ with smooth Lagrangian Ł

$$\int Ł(t,y_1,\ldots,y_n,y_1',\ldots,y_n')dt \qquad \left(y_i = y_i(t),\ ' = \frac{d}{dt},\ \det\left(\frac{\partial^2 Ł}{\partial y_i' \partial y_j'}\right) \neq 0\right) \tag{8.3}$$

to appear later, involves variables from quite other jet space $\mathbf{M}(n,1)$ with coordinates denoted t (the independent variable), $y_1,\ldots,y_n$ (the dependent variables) and higher-order jet variables like y_i', y_i'' and so on.

We are passing to the topic proper. Let us start in the space $\mathbf{M}(1,n)$ with $\mathcal{CT}$ groups. One can check that vector field (5.6) is infinitesimal $\mathcal{CT}$ if and only if

$$Z = -\sum Q_{w_i}\frac{\partial}{\partial x_i} + \left(Q - \sum w_i Q_{w_i}\right)\frac{\partial}{\partial w} + \sum (Q_{x_i} + w_i Q_w)\frac{\partial}{\partial w_i} + \cdots, \tag{8.4}$$

where the function $Q = Q(x_1,\ldots,x_n,w,w_1,\ldots,w_n)$ may be arbitrarily chosen.

"Hint: we have, by definition

$$\mathcal{L}_Z \omega = Z \rfloor \mathrm{d}\omega + \mathrm{d}\omega(Z) = \sum (z_i \omega_i - \omega_i(Z)\mathrm{d}x_i) + \mathrm{d}Q \in \Omega_0, \tag{8.5}$$

where $Q = Q(x_1,\ldots,x_n,w,w_1,\ldots,w_n,\ldots) = \omega(Z) = z^1 - \sum w_i z_i$,

$$\mathrm{d}Q = \sum D_i Q \mathrm{d}x_i + \frac{\partial Q}{\partial w}\omega + \sum \frac{\partial Q}{\partial w_i}\omega_i \tag{8.6}$$

to Contents, page 1

whence immediately $z_i = -\partial Q/\partial w_i$, $z^1 = Q + \sum w_i z_i = Q - \sum w_i \cdot \partial Q/\partial w_i$, $\partial Q/\partial w_I = 0$ if $|I| \geq 1$ and formula (8.4) follows."

Alas, the corresponding Lie system (not written here) is not much inspirational. Let us however consider a function $w = w(x_1, \ldots, x_n)$ implicitly defined by an equation $V(x_1, \ldots, x_n, w) = 0$. We may suppose that the transformed function $\mathbf{m}(\lambda)^* w$ satisfies the equation

$$V(x_1, \ldots, x_n, \mathbf{m}(\lambda)^* w) = \lambda \tag{8.7}$$

without any loss of generality. In infinitesimal terms

$$1 = \frac{\partial(V - \lambda)}{\partial \lambda} = Z(V - \lambda) = -\sum Q_{w_i} V_{x_i} + \left(Q - \sum w_i Q_{w_i}\right) V_w. \tag{8.8}$$

However $w_i = \partial w/\partial x_i = -V_{x_i}/V_w$ may be inserted here, and we have the crucial *Jacobi equation*

$$1 = Q\left(x_1, \ldots, x_n, w, -\frac{V_{x_1}}{V_w}, \ldots, -\frac{V_{x_n}}{V_w}\right) V_w \tag{8.9}$$

(not involving V) which can be uniquely rewritten as the *Hamilton-Jacobi* ($\mathcal{HJ}$) *equation*

$$V_w + \mathcal{H}(x_1, \ldots, x_n, w, p_1, \ldots, p_n) \quad (p_i = V_{x_i}) \tag{8.10}$$

in the "nondegenerate" case $\sum Q_{w_i} V_{x_i} \neq 1$. Let us recall the characteristic curves [22, 23] of the $\mathcal{HJ}$ equation given by the system

$$\frac{\mathrm{d}w}{1} = \frac{\mathrm{d}x_i}{\mathcal{H}_{p_i}} = -\frac{\mathrm{d}p_i}{\mathcal{H}_{x_i}} = \frac{\mathrm{d}V}{-\mathcal{H} + \sum p_i \mathcal{H}_{p_i}}. \tag{8.11}$$

The curves may be interpreted as the orbits of the group $\mathbf{m}(\lambda)$. (Hint: look at the well-known classical construction of the solution V of the Cauchy problem [22, 23] in terms of the characteristics. The initial Cauchy data are transferred just along the characteristics, i.e., along the group orbits.) Assume moreover the additional condition $\det(\partial^2\mathcal{H}/\partial p_i \partial p_j) \neq 0$. We may introduce variational integral (8.3) with the Lagrange function Ł given by the familiar identities

$$Ł + \mathcal{H} = \sum p_i y_i' \tag{8.12}$$

with interrelations

$$t = w, \quad y_i = x_i, \quad y_i' = \mathcal{H}_{p_i}, \quad p_i = Ł_{y_i'} \quad (i = 1, \ldots, n) \tag{8.13}$$

between variables t, y_i, y_i' of the space $\mathbf{M}(n, 1)$ and variables x_i, w, w_i of the space $\mathbf{M}(1, n)$. Since (8.11) may be regarded as a Hamiltonian system for the extremals of $\mathcal{VJ}$, the metatheorem is clarified. □

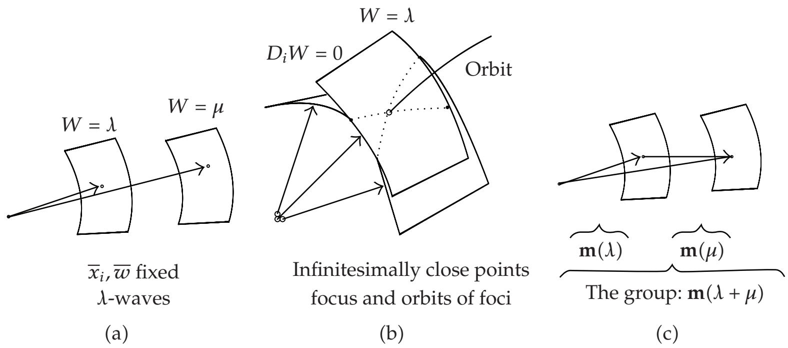


Figure 3

Remark 8.2. Let us recall the Mayer fields of extremals for the $\mathcal{VJ}$ since they provide the true sense of the above construction. The familiar Poincaré-Cartan form

$$\breve{\varphi} = Ł\mathrm{d}t + \sum Ł_{y_i'}\left(\mathrm{d}y_i - y_i'\mathrm{d}t\right) = -\mathscr{H}\mathrm{d}t + \sum p_i\mathrm{d}y_i \tag{8.14}$$

is restricted to appropriate subspace $y_i' = g_i(t, y_1, \ldots, y_n)$ $(i = 1, \ldots, n$; the *slope field*) in order to become a total differential

$$\breve{\varphi}\big|_{y_i'=g_i} = \mathrm{d}V(t, y_1, \ldots, y_n) = V_t\mathrm{d}t + \sum V_{y_i}\mathrm{d}y_i \tag{8.15}$$

of the *action* V. We obtain the requirements $V_t = -\mathscr{H}$, $V_{y_i} = p_i$ identical with (8.10). In geometrical terms: *transformations of a hypersurface* $V = 0$ *by means of* $\mathcal{CT}$ *group may be identified with the level sets* $V = \lambda$ $(\lambda \in \mathbb{R})$ *of the action of a Mayer fields of extremals.*

The last statement is in accordance with (8.11) where

$$\mathrm{d}V = \left(-\mathscr{H} + \sum p_i\mathscr{H}_{p_i}\right)\mathrm{d}w = \left(-\mathscr{H} + \sum p_i y_i'\right)\mathrm{d}t = Ł\mathrm{d}t, \tag{8.16}$$

use the identifications (8.13) of coordinates. This is the classical definition of the action V in a Mayer field. We have moreover clarified the additive nature of the level sets $V = \lambda$: roughly saying, the composition with $V = \mu$ provides $V = \lambda + \mu$ (see Figure 3(c)) and this is caused by the additivity of the integral $\int Ł\ dt$ calculated along the orbits.

On this occasion, the wave enveloping approach to $\mathcal{CT}$ groups is also worth mentioning.

Lemma 8.3 (see [10–13]). *Let* $W(\overline{x}_1, \ldots, \overline{x}_n, \overline{w}, x_1, \ldots, x_n, w)$ *be a function of* $2n + 2$ *variables. Assume that the system* $W = D_1W = \cdots = D_nW = 0$ *admits a unique solution*

$$\overline{x}_i = F_i(\ldots, x_i{}', w, w_i{}', \ldots), \qquad \overline{w} = F^1(\ldots, x_i{}', w, w_i{}', \ldots) \tag{8.17}$$

to Contents, page 1

by applying the implicit function theorem and analogously the system $W = \overline{D}_1 W = \cdots = \overline{D}_n W = 0$ *(where* $\overline{D}_i = \partial/\partial\overline{x}_i + \sum \overline{w}_i \partial/\partial\overline{w}$*) admits a certain solution*

$$x_i = \overline{F}_i(\ldots, \overline{x}_{i'}, \overline{w}, \overline{w}_{i'}, \ldots), \qquad w = \overline{F}^1(\ldots, \overline{x}_{i'}, \overline{w}, \overline{w}_{i'}, \ldots). \tag{8.18}$$

Then $\mathbf{m}^* x_i = F_i$, $\mathbf{m}^* w = F^1$ *provides a Lie* $\mathcal{CT}$ *and* $(\mathbf{m}^{-1})^* \overline{x}_i = \overline{F}_i$, $(\mathbf{m}^{-1})^* \overline{w} = \overline{F}^1$ *is the inverse.*

In more generality, if function W in Lemma 8.3 moreover depends on a parameter λ, we obtain a mapping $\mathbf{m}(\lambda)$ which is a certain $\mathcal{CT}$ involving a parameter λ and the inverse $\mathbf{m}(\lambda)^{-1}$. *In favourable case* (see below) *this* $\mathbf{m}(\lambda)$ *may be even a* $\mathcal{CT}$ *group*. The geometrical sense is as follows. Equation $W = 0$ with $\overline{x}_i$, $\overline{w}$ kept fixed represents a wave in the space x_i, w (Figure 3(a)).

The total system $W = D_1 W = \cdots = D_n W = 0$ provides the intersection (envelope) of infinitely close waves (Figure 3(b)) with the resulting transform, the focus point $\mathbf{m}$ (or $\mathbf{m}(\lambda)$ if the parameter λ is present). The reverse waves with the role of variables interchanged gives the inversion. Then the group property holds true if the waves can be composed (Figure 3(c)) within the parameters λ, μ, but this *need not* be in general the case.

Let us eventually deal with the condition ensuring the group composition property. Without loss of generality, we may consider the λ-depending wave

$$W(\overline{x}_1, \ldots, \overline{x}_n, \overline{w}, x_1, \ldots, x_n, w) - \lambda = 0. \tag{8.19}$$

If $\overline{x}_i$, $\overline{w}$ are kept fixed, the previous results may be applied. We obtain a group if and only if the $\mathcal{HJ}$ equation (8.10) holds true, therefore

$$W_w + \mathcal{H}(x_1, \ldots, x_n, w, W_{x_1}, \ldots, W_{x_n}) = 0. \tag{8.20}$$

The existence of such function $\mathcal{H}$ means that functions $W_w, W_{x_1}, \ldots, W_{x_n}$ *of dashed variables* are functionally dependent whence

$$\det\begin{pmatrix} W_{w\overline{w}} & W_{w\overline{x}_{i'}} \\ W_{x_i\overline{w}} & W_{x_i\overline{x}_{i'}} \end{pmatrix} = 0, \qquad \det\left(W_{x_i\overline{x}_{i'}}\right) \neq 0. \tag{8.21}$$

The symmetry $\overline{x}_i, \overline{w} \leftrightarrow x_i, w$ is not surprising here since the change $\lambda \leftrightarrow -\lambda$ provides the inverse mapping: equations

$$W(\ldots, \overline{x}_i, \overline{w}, \ldots, x_i, w) = \lambda, \qquad W(\ldots, x_i, w, \ldots, \overline{x}_i, \overline{w}) = -\lambda \tag{8.22}$$

are equivalent. In particular, it follows that

$$W(\ldots, \overline{x}_i, \overline{w}, \ldots, x_i, w) = -W(\ldots, x_i, w, \ldots, \overline{x}_i, \overline{w}), \qquad W(\ldots, x_i, w, \ldots, x_i, w) = 0 \tag{8.23}$$

and the wave $W - \lambda = 0$ corresponds to the Mayer central field of extremals.

to Contents, page 1

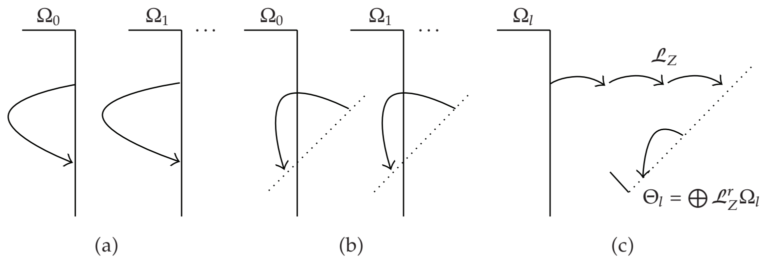


Figure 4

Summary 4. Conditions (8.21) ensure the existence of $\mathcal{H}\mathcal{J}$ equation (8.20) for the λ-wave (8.19) and therefore the group composition property of waves (8.19) in the nondegenerate case $\det(\partial^2\mathcal{H}/\partial p_i \partial p_j) \neq 0$.

Remark 8.4. A reasonable theory of Mayer fields of extremals and Hamilton-Jacobi equations can be developed also for the constrained variational integrals (the Lagrange problem) within the framework of jet spaces, that is, without the additional Lagrange multipliers [9, Chapter 3]. It follows that there do exist certain groups of generalized Lie's contact transformations with differential constraints.

9. On the Order-Destroying Groups in Jet Space

We recall that in the order-preserving case, the filtration

$$\Omega(m,n)_* : \Omega_0 \subset \Omega_1 \subset \cdots \subset \Omega(m,n) = \cup \Omega_l \tag{9.1}$$

of module $\Omega(m,n)$ is preserved (Figure 4(a)). It follows that certain invariant submodules $\Omega_l \subset \Omega(m,n)$ are a priori prescribed which essentially restricts the store of the symmetries (the Lie-Bäcklund theorem). The order-destroying groups also preserve certain submodules of $\Omega(m,n)$ due to approximation results, however, they are not known in advance (Figure 4(b)) and appear after certain saturation (Figure 4(c)) described in technical theorem 5.1.

The saturation is in general a toilsome procedure. It may be simplified by applying two simple principles.

Lemma 9.1 (going-up lemma). *Let a group of morphisms* $\mathbf{m}(\lambda)$ *preserve a submodule* $\Theta \subset \Omega(m,n)$. *Then also the submodule*

$$\Theta + \sum \mathcal{L}_{D_i}\Theta \subset \Omega(m,n) \tag{9.2}$$

is preserved.

Proof. We suppose $\mathcal{L}_Z\Theta \subset \Theta$. Then

$$\mathcal{L}_Z\left(\Theta + \sum \mathcal{L}_{D_i}\Theta\right) = \mathcal{L}_Z\Theta + \left(\mathcal{L}_{D_i}\mathcal{L}_Z\Theta - \sum D_i z_{i'} \mathcal{L}_{D_{i'}}\Theta\right) \subset \Theta + \sum \mathcal{L}_{D_i}\Theta \tag{9.3}$$

by using the commutative rule (5.17). ☐

to Contents, page 1

Lemma 9.2 (going-down lemma). *Let the group of morphisms* $\mathbf{m}(\lambda)$ *preserve a submodule* $\Theta \subset \Omega(m,n)$. *Let* $\Theta' \subset \Theta$ *be the submodule of all* $\omega \in \Theta$ *satisfying* $\mathcal{L}_{D_i}\omega \in \Theta$ $(i = 1,\dots,n)$. *Then* Θ' *is preserved, too.*

Proof. Assume $\omega \in \Theta'$ hence $\mathcal{L}_{D_i}\omega \in \Theta$. Then $\mathcal{L}_{D_i}\mathcal{L}_Z\omega = \mathcal{L}_Z\mathcal{L}_{D_i}\omega + \mathcal{L}_{\sum D_i z_{i'} \cdot D_{i'}}\omega \in \Theta$ hence $\mathcal{L}_Z\omega \in \Theta'$ and Θ' is preserved. □

We are passing to illustrative examples.

Example 9.3. Let us consider the vector field (the variation of jet structure)

$$Z = \sum z_I^j \frac{\partial}{\partial w_I^j} \quad \left(z_I^j = D_I z^j,\ D_I = D_{i_1} \cdots D_{i_n}\right), \tag{9.4}$$

see (5.6) and (5.10) for the particular case $z_i = 0$. Then $Z^r x_i = 0$ $(i = 1,\dots,n)$ and the *sufficient* requirement $Z^2 w^j = 0$ $(j = 1,\dots,m)$ ensures $Z \in \mathbb{G}$, see (ι) of Lemma 5.4. We will deal with the linear case where

$$z^j = \sum a_{i'}^{jj'} w_{i'}^{j'} \quad \left(a_{i'}^{jj'} \in \mathbb{R}\right) \tag{9.5}$$

is supposed. Then

$$Z^2 w^j = Z z^j = \sum a_{i'}^{jj'} z_{i'}^{j'} = \sum a_{i'}^{jj'} a_{i''}^{j'j''} w_{i'i''}^{j''} = 0 \tag{9.6}$$

identically if and only if

$$\sum_{j'} \left(a_{i'}^{jj'} a_{i''}^{j'j''} + a_{i''}^{jj'} a_{i'}^{j'j''}\right) = 0 \quad \left(i', i'' = 1,\dots,n;\ j, j', j'' = 1,\dots,m\right). \tag{9.7}$$

This may be expressed in terms of matrix equations

$$A_i A_{i'} = 0 \quad \left(i, i' = 1,\dots,n;\ A_i = \left(a_i^{jj'}\right)\right) \tag{9.8}$$

or, in either of more geometrical transcriptions

$$A^2 = 0, \qquad \operatorname{Im} A \subset \operatorname{Ker} A \quad \left(A = \sum \lambda_i A_i,\ \lambda_i \in \mathbb{R}\right), \tag{9.9}$$

where A is regarded as (a matrix of an) operator acting in m-dimensional linear space and depending on parameters $\lambda_1,\dots,\lambda_n$. We do not know explicit solutions A in full generality, however, solutions A such that Ker A does not depend on the parameters $\lambda_1,\dots,\lambda_n$ can be easily found (and need not be stated here). The same approach can be applied to the more general *sufficient* requirement $Z^r w^j = 0$ $(j = 1,\dots,m$; fixed $r)$ ensuring $Z \in \mathbb{G}$. If $r \geq n$, the requirement is *equivalent* to the inclusion $Z \in \mathbb{G}$.

to Contents, page 1

Example 9.4. Let us consider vector field (5.6) where $z^1 = \cdots = z^m = 0$. In more detail, we take

$$Z = \sum z_i \frac{\partial}{\partial x_i} + \sum z_i^j \frac{\partial}{\partial w_i^j} + \cdots \quad \left(z_i^j = -\sum w_{i'}^j D_i z_{i'} \right). \tag{9.10}$$

Then $Z^r w^j = 0$ and we have to deal with functions $Z^r x_i$ in order to ensure the inclusion $Z \in \mathbb{G}$. This is a difficult task. Let us therefore suppose

$$z_1 = z\left(\ldots, x_{i'}, w^{j'}, w_1^{j'}, \ldots \right), \qquad z_k = c_k \in \mathbb{R} \quad (k = 2, \ldots, n). \tag{9.11}$$

Then $Zx_k = 0$ $(k = 2, \ldots, n)$ and

$$Z^2 x_1 = Zz = \sum \frac{\partial z}{\partial x_i} z_i + \sum \frac{\partial z}{\partial w_1^j} z_1^j, \tag{9.12}$$

where

$$z_1^j = -w_1^j D_1 z = -w_1^j \left(\frac{\partial z}{\partial x_1} + \sum \frac{\partial z}{\partial w^{j'}} w_1^{j'} + \sum \frac{\partial z}{\partial w_1^{j'}} w_{11}^{j'} \right). \tag{9.13}$$

The second-order summand

$$Z^2 x_1 = \cdots + \sum \frac{\partial z}{\partial w_1^j} z_1^j = \cdots - \sum \frac{\partial z}{\partial w_1^j} w_1^j \frac{\partial z}{\partial w_1^{j'}} w_{11}^{j'} \tag{9.14}$$

identically vanishes for the choice

$$z = f\left(\ldots, x_{i'}, w^{j'}, u^l, \ldots \right) \quad \left(u^l = \frac{w_1^l}{w_1^1};\ l = 2, \ldots, m \right) \tag{9.15}$$

as follows by direct verification. Quite analogously

$$Zu^l = Z\frac{w_1^l}{w_1^1} = z_1^l \frac{1}{w_1^1} - z_1^1 \frac{w_1^l}{\left(w_1^1\right)^2} = \left(-w_1^l \frac{1}{w_1^1} + w_1^1 \frac{w_1^l}{\left(w_1^1\right)^2} \right) D_1 z = 0. \tag{9.16}$$

It follows that all functions $Z^r x_i$, $Z^r w^j$ can be expressed in terms of the *finite family* of functions x_i $(i = 1, \ldots, n)$, w^j $(j = 1, \ldots, m)$, u^l $(l = 2, \ldots, m)$ and therefore $Z \in \mathbb{G}$.

Remark 9.5. On this occasion, let us briefly mention the groups generated by vector fields Z of the above examples. The Lie system of the vector field (9.4) and (9.5) reads

$$\frac{\mathrm{d}G_i}{\mathrm{d}\lambda} = 0, \qquad \frac{\mathrm{d}G^j}{\mathrm{d}\lambda} = \sum a_{i'}^{jj'} G_{i'}^{j'} \quad (i = 1, \ldots, n;\ j = 1, \ldots, m), \tag{9.17}$$

to Contents, page 1

where we omit the prolongations. It is resolved by

$$G_i = x_i, \quad G^j = w^j + \lambda \sum a^{jj'}_{i'} w^{j'}_{i'} \quad (i = 1, \ldots, n;\ j = 1, \ldots, m) \tag{9.18}$$

as follows either by direct verification or, alternatively, from the property $Z^2 x_i = Zz_i = 0$ ($i = 1, \ldots, n$) which implies

$$\frac{\mathrm{d} \sum a^{jj'}_{i'} G^{j'}_{i'}}{\mathrm{d}\lambda} = 0, \qquad \sum a^{jj'}_{i'} G^{j'}_{i'} = \sum a^{jj'}_{i'} G^{j'}_{i'} \Big|_{\lambda=0} = \sum a^{jj'}_{i'} w^{j'}_{i'}. \tag{9.19}$$

Quite analogously, the Lie system of the vector field (9.10), (9.11), (9.15) reads

$$\frac{\mathrm{d}G_1}{\mathrm{d}\lambda} = f\left(\ldots, G_{i'}, G^{j'}, \frac{G^{l'}_1}{G^1_1}, \ldots\right), \quad \frac{\mathrm{d}G_k}{\mathrm{d}\lambda} = c_k, \quad \frac{\mathrm{d}G^j}{\mathrm{d}\lambda} = 0 \quad (k = 2, \ldots, n;\ j = 1, \ldots, m) \tag{9.20}$$

and may be completed with the equations

$$\frac{\mathrm{d}\left(G^l_1 / G^1_1\right)}{\mathrm{d}\lambda} = 0 \quad (l = 2, \ldots, m) \tag{9.21}$$

following from (9.16). This provides a classical self-contained system of ordinary differential equations where the common existence theorems can be applied.

The above Lie systems admit many nontrivial *first integrals* $F \in \mathcal{F}$, that is, functions F that are constant on the orbits of the group. Conditions $F = 0$ may be interpreted as differential equations in the total jet space, and the above transformation groups turn into the *external generalized symmetries* of such differential equations, see Section 11 below.

10. Towards the Main Algorithm

We briefly recall the algorithm [10–13] for determination of all individual automorphisms $\mathbf{m}$ of the jet space $\mathbf{M}(m, n)$ in order to compare it with the subsequent calculation of vector field $Z \in \mathbb{G}$.

Morphisms $\mathbf{m}$ of the jet structure were defined by the property $\mathbf{m}^* \Omega(m, n) \subset \Omega(m, n)$. The inverse $\mathbf{m}^{-1}$ exists if and only if

$$\Omega_0 \subset \mathbf{m}^* \Omega(m, n), \quad \text{equivalently } \Omega_0 \subset \mathbf{m}^* \Omega_l \quad (l = l(\mathbf{m})) \tag{10.1}$$

for appropriate term $\Omega_{l(\mathbf{m})}$ of filtration (9.1). However

$$\mathbf{m}^* \Omega_{l+1} = \mathbf{m}^* \Omega_l + \sum \mathcal{L}_{D_i} \mathbf{m}^* \Omega_l \tag{10.2}$$

to Contents, page 1

and it follows that criterion (10.1) can be verified by repeated use of operators $\mathcal{L}_{D_i}$. In more detail, we start with equations

$$\mathbf{m}^*\omega^j = \sum a^{jj'}_{I'}\omega^{j'}_{I'} \quad \left(= \mathrm{d}\mathbf{m}^*w^j - \sum \mathbf{m}^*w^j_i \mathrm{d}\mathbf{m}^*x_i\right) \tag{10.3}$$

with uncertain coefficients. Formulae (10.3) determine the module $\mathbf{m}^*\Omega_0$. Then we search for lower-order contact forms, especially forms from Ω_0, lying in $\mathbf{m}^*\Omega_l$ with the use of (10.2). Such forms are ensured if certain *linear relations among coefficients exist*. The calculation is finished on a certain level $l = l(\mathbf{m})$ and this is the *algebraic part* of the algorithm. With this favourable choice of coefficients $a^{jj'}_{I'}$, functions $\mathbf{m}^*x_i$, $\mathbf{m}^*w^j$ (and therefore the invertible morphism $\mathbf{m}$) can be determined by inspection of the bracket in (10.3). This is the *analytic part* of algorithm.

Let us turn to the infinitesimal theory. Then the main technical tool is the rule (5.17) in the following transcription:

$$\mathcal{L}_Z\mathcal{L}_{D_i} = \mathcal{L}_{D_i}\mathcal{L}_Z - \sum D_i z_{i'}\mathcal{L}_{D_{i'}} \tag{10.4}$$

or, when applied to basical forms

$$\mathcal{L}_Z\omega^j_{Ii} = \mathcal{L}_{D_i}\mathcal{L}_Z\omega^j_I - \sum D_i z_{i'}\omega^j_{Ii'}. \tag{10.5}$$

We are interested in vector fields $Z \in \mathbb{G}$. They satisfy the recurrence (5.10) together with requirements

$$\dim \oplus \mathcal{L}^r_Z\Omega_0 < \infty, \quad \text{equivalently } \mathcal{L}^r_Z\Omega_0 \subset \Omega_{l(Z)} \quad (r = 0, 1, \ldots) \tag{10.6}$$

for appropriate $l(Z) \in \mathbb{N}$. Due to the recurrence (10.5) these requirements can be effectively investigated. In more detail, we start with equations

$$\mathcal{L}_Z\omega^j = \sum a^{jj'}_{I'}\omega^{j'}_{I'} \quad \left(= \mathrm{d}z^j - \sum z^j_i \mathrm{d}x_i - \sum w^j_i \mathrm{d}z_i\right). \tag{10.7}$$

Formulae (10.7) determine module $\mathcal{L}_Z\Omega_0$. Then, choosing $l(Z) \in \mathbb{N}$, operator $\mathcal{L}_Z$ is to be repeatedly applied and requirements (10.6) provide certain *polynomial relations for the coefficients* by using (10.5). This is the *algebraical part* of the algorithm. With such coefficients $a^{jj'}_{I'}$ available, functions $z_i = \mathcal{L}_Z x_i$, $z^j = \mathcal{L}_Z w^j$ (and therefore the vector field $Z \in \mathbb{G}$) can be determined by inspection of the bracket in (10.7) or, alternatively, with the use of formulae (5.12) for the particular case $I = \phi$ empty

$$\mathcal{L}_Z\omega^j = \sum \left(\frac{\partial z^j}{\partial w^{j'}_{I'}} - \sum w^j_i \frac{\partial z_i}{\partial w^{j'}_{I'}} \right)\omega^{j'}_{I'}. \tag{10.8}$$

This is the *analytic part* of the algorithm.

to Contents, page 1

Altogether taken, the algorithm is not easy and the conviction [7, page 121] that the "exhaustive description of integrable C-fields (fields $Z \in \mathbb{Z}$ in our notation) is given in [16]" is disputable. We can state only one optimistic result at this place.

Theorem 10.1. *The jet spaces* $\mathbf{M}(1,n)$ *do not admit any true generalized infinitesimal symmetries* $Z \in \mathbb{G}$.

Proof. We suppose $m = 1$ and then (10.7) reads

$$\mathcal{L}_Z \omega^1 = \sum a^{11}_{I'} \omega^1_{I'} = \cdots + a^{11}_{I''} \omega^1_{I''} \quad \left(a^{11}_{I''} \neq 0\right), \tag{10.9}$$

where we state a summand of maximal order. Assuming $I'' = \phi$, the Lie-Bäcklund theorem can be applied and we do not have the true generalized symmetry Z. Assuming $I'' \neq \phi$, then

$$\mathcal{L}^r_Z \omega^1 = \cdots + a^{11}_{I''} \omega^1_{I'' \ldots I''} \quad (r \text{ terms } I'') \tag{10.10}$$

by using rule (10.5) where the last summand may be omitted. It follows that (10.6) is not satisfied hence $Z \notin \mathbb{G}$. □

Example 10.2. We discuss the simplest possible but still a nontrivial particular example. Assume $m = 2$, $n = 1$ and $l(Z) = 1$. Let us abbreviate

$$x = x_1, \quad D = D_1, \quad Z = z\frac{\partial}{\partial x} + \sum z^j_I \frac{\partial}{\partial w^j_I} \quad (j = 1, 2;\ I = 1 \cdots 1). \tag{10.11}$$

Then, due to $l(Z) = 1$, requirement (10.6) reads

$$\mathcal{L}^r_Z \Omega_0 \subset \Omega_1 \quad (r = 0, 1, \ldots). \tag{10.12}$$

In particular (if $r = 1$) we have (10.7) written here in the simplified notation

$$\mathcal{L}_Z \omega^j = a^{j1}\omega^1 + a^{j2}\omega^2 + b^{j1}\omega^1_1 + b^{j2}\omega^2_1 \quad (j = 1, 2). \tag{10.13}$$

The next requirement ($r = 2$) *implies the* (only seemingly) *stronger inclusion*

$$\mathcal{L}^2_Z \Omega_0 \subset \mathcal{L}_Z \Omega_0 + \Omega_0 \tag{10.14}$$

which already ensures (10.12) for all r and therefore $Z \in \mathbb{G}$ (easy). *We suppose* (10.14) *from now on*.

"Hint for proof of (10.14): assuming (10.12) and moreover the equality

$$\mathcal{L}^2_Z \Omega_0 + \mathcal{L}_Z \Omega_0 + \Omega_0 = \Omega_1, \tag{10.15}$$

to Contents, page 1

it follows that

$$\mathcal{L}_Z\Omega_1 \subset \mathcal{L}_Z^3\Omega_0 + \mathcal{L}_Z^2\Omega_0 + \mathcal{L}_Z\Omega_0 \subset \Omega_1 \tag{10.16}$$

and Lie-Bäcklund theorem can be applied whence $\mathcal{L}_Z\Omega_0 \subset \Omega_0$, $l(Z) = 0$ which we exclude. It follows that necessarily

$$\dim\left(\mathcal{L}_Z^2\Omega_0 + \mathcal{L}_Z\Omega_0 + \Omega_0\right) < \dim\Omega_1 = 4. \tag{10.17}$$

On the other hand $\dim(\mathcal{L}_Z\Omega_0 + \Omega_0) \geq 3$ and the inclusion (10.14) follows."

After this preparation, we are passing to the proper algebra. Clearly

$$\mathcal{L}_Z^2\omega^j = \cdots + b^{j1}\mathcal{L}_Z\omega_1^1 + b^{j2}\mathcal{L}_Z\omega_1^2 = \cdots + b^{j1}\left(b^{11}\omega_{11}^1 + b^{12}\omega_{11}^2\right) + b^{j2}\left(b^{21}\omega_{11}^1 + b^{22}\omega_{11}^2\right) \tag{10.18}$$

by using the commutative rule (10.5). Due to "weaker" inclusion (10.12) with $r = 2$, we obtain identities

$$b^{j1}b^{11} + b^{j2}b^{21} = 0, \quad b^{j1}b^{12} + b^{j2}b^{22} = 0 \quad (j = 1, 2). \tag{10.19}$$

Omitting the trivial solution, they are satisfied if either

$$b^{11} + b^{22} = 0, \qquad b^{12} = cb^{11}, \qquad b^{11} + cb^{21} = 0 \tag{10.20}$$

for appropriate factor c (where $b^{11} \neq 0$ and either $b^{12} \neq 0$ or $b^{21} \neq 0$ is supposed) or

$$b^{11} = b^{22} = 0, \quad \text{either } b^{12} = 0 \quad \text{or } b^{21} = 0. \tag{10.21}$$

We deal only with the (more interesting) identities (10.20) here. Then

$$\begin{aligned} \mathcal{L}_Z\omega^1 &= a^{11}\omega^1 + a^{12}\omega^2 - cb\left(\omega_1^1 + c\omega_1^2\right), \\ \mathcal{L}_Z\omega^2 &= a^{21}\omega^1 + a^{22}\omega^2 + b\left(\omega_1^1 + c\omega_1^2\right) \end{aligned} \tag{10.22}$$

(abbreviation $b = b^{21}$) by inserting (10.20) into (10.13). It follows that

$$\mathcal{L}_Z\left(\omega^1 + c\omega^2\right) = a^1\omega^1 + a^2\omega^2 \quad \left(a^1 = a^{11} + ca^{21}, a^2 = a^{12} + ca^{22} + Zc\right). \tag{10.23}$$

It may be seen by direct calculation of $\mathcal{L}_Z^2\omega^2$ that the "stronger" inclusion (10.14) is equivalent

to Contents, page 1

to the identity $ca^1 = a^2$, that is,

$$\mathcal{L}_Z\left(\omega^1 + c\omega^2\right) = a\left(\omega^1 + c\omega^2\right) \tag{10.24}$$

(abbreviation $a = a^1$). Alternatively, (10.24) can be proved by using Lemma 9.2.

"Hint: denoting $\Theta = \mathcal{L}_Z\Omega_0 + \Omega_0$, (10.14) implies $\mathcal{L}_Z\Theta \subset \Theta$. Moreover $\mathcal{L}_D(\omega^1 + c\omega^2) \in \Theta$ by using (10.22). Lemma 9.2 can be applied: $\omega^1 + c\omega^2 \in \Theta'$ and Θ' involves just all multiples of form $\omega^1 + c\omega^2$. Therefore $\mathcal{L}_Z(\omega^1 + c\omega^2) \in \Theta'$ is a multiple of $\omega^1 + c\omega^2$."

The algebraical part is concluded. We have congruences

$$\mathcal{L}_Z\omega^1 \cong -cb\left(\omega_1^1 + c\omega_1^2\right), \quad \mathcal{L}_Z\omega^2 \cong b\left(\omega_1^1 + c\omega_1^2\right) \quad (\bmod\ \Omega_0) \tag{10.25}$$

and equality

$$\mathcal{L}_Z\omega^1 + c\mathcal{L}_Z\omega^2 + Zc\omega^2 = a\left(\omega^1 + c\omega^2\right). \tag{10.26}$$

If Z is a variation then these three conditions together ensure the "stronger inclusion" (10.14) hence $Z \in \mathbb{G}$.

We turn to analysis. Abbreviating

$$Z_{I'}^{jj'} = \frac{\partial z^j}{\partial w_{I'}^{j'}} - w_1^j \frac{\partial z}{\partial w_{I'}^{j'}} \quad \left(j, j' = 1, 2;\ I' = 1 \cdots 1\right) \tag{10.27}$$

and employing (10.8), the above conditions (10.25) and (10.26) read

$$\begin{gathered}\sum Z_{I'}^{1j'}\omega_{I'}^{j'} = -cb\left(\omega_1^1 + c\omega_1^2\right), \quad \sum Z_{I'}^{2j'}\omega_{I'}^{j'} = b\left(\omega_1^1 + c\omega_1^2\right) \quad \left(\left|I'\right| \geq 1\right), \\ \sum\left(Z_{I'}^{1j'} + cZ_{I'}^{2j'}\right)\omega_{I'}^{j'} + Zc\omega^2 = a\left(\omega^1 + c\omega^2\right).\end{gathered} \tag{10.28}$$

We compare coefficients of forms ω_I^j on the level $s = |I'|$

$$s = 0\colon\ Z^{11} + cZ^{21} = a, \qquad Z^{12} + cZ^{22} + Zc = ac, \tag{10.29}$$

$$s = 1\colon\ Z_1^{11} = -cb, \qquad Z_1^{12} = -(c)^2 b, \qquad Z_1^{21} = b, \qquad Z_1^{22} = bc, \qquad Z_1^{1j} + cZ_1^{2j} = 0, \tag{10.30}$$

$$s \geq 2\colon\ Z_{I'}^{jj'} = 0, \qquad Z_{I'}^{1j'} + cZ_{I'}^{2j'} = 0. \tag{10.31}$$

We will successively delete the coefficients a, b, c in order to obtain interrelations only for variables $Z_{I'}^{jj'}$. Clearly

$$\begin{gathered}s = 0\colon\ Z^{12} + cZ^{22} + Zc = \left(Z^{11} + cZ^{21}\right)c, \\ s = 1\colon\ Z_1^{11} + Z_1^{22} = 0, \qquad Z_1^{11}Z_1^{22} = Z_1^{12}Z_1^{21},\end{gathered} \tag{10.32}$$

to Contents, page 1

and we moreover have three compatible equations

$$c = -\frac{Z_1^{11}}{Z_1^{21}} = -\frac{Z_1^{12}}{Z_1^{22}}, \qquad (c)^2 = -\frac{Z_1^{12}}{Z_1^{21}} \tag{10.33}$$

for the coefficient c. To cope with levels $s \geq 2$, we introduce functions

$$Q^j = \omega^j(Z) = z^j - w_1^j z \quad (j = 1, 2). \tag{10.34}$$

Then substitution into (10.27) with the help of (10.31) gives

$$\frac{\partial Q^j}{\partial w_{I'}^{j'}} = 0 \quad \left(j, j' = 1, 2;\ |I'| \geq 2\right). \tag{10.35}$$

It follows moreover easily that

$$Z_1^{jj'} = \frac{\partial Q^j}{\partial w_1^{j'}} \quad (j \neq j'), \qquad Z_1^{jj} = z + \frac{\partial Q^j}{\partial w_1^j}, \qquad Z^{jj'} = \frac{\partial Q^j}{\partial w^{j'}} \tag{10.36}$$

and we have *the final differential equations*

$$s = 0\colon\ \frac{\partial Q^1}{\partial w^2} + c\frac{\partial Q^2}{\partial w^2} + Zc = \left(\frac{\partial Q^1}{\partial w^1} + c\frac{\partial Q^2}{\partial w^1}\right)c, \tag{10.37}$$

$$s = 1\colon 2z + \frac{\partial Q^1}{\partial w_1^1} + \frac{\partial Q^2}{\partial w_1^2} = 0, \quad \left(z + \frac{\partial Q^1}{\partial w_1^1}\right)\left(z + \frac{\partial Q^2}{\partial w_1^2}\right) = \frac{\partial Q^1}{\partial w_1^2}\frac{\partial Q^2}{\partial w_1^1} \tag{10.38}$$

for the unknown functions

$$z = z\left(x, w^1, w^2, w_1^1, w_1^2\right), \qquad Q^j = Q^j\left(x, w^1, w^2, w_1^1, w_1^2\right). \tag{10.39}$$

The coefficient c is determined by (10.33) *and* (10.36) *in terms of functions Q^j. This concludes the analytic part of the algorithm since trivially $z^j = w_1^j z + Q^j$ and the vector field Z is determined.*

The system is compatible: particular solutions with functions Q^j quadratic in jet variables and $c = \text{const.}$ can be found as follows. Assume

$$Q^j = A^j\left(w_1^1\right)^2 + 2B^j w_1^1 w_1^2 + C^j\left(w_1^2\right)^2 \quad (j = 1, 2) \tag{10.40}$$

with constant coefficients $A^j, B^j, C^j \in \mathbb{R}$. We also suppose $c \in \mathbb{R}$ and then (10.37) is trivially satisfied.

to Contents, page 1

On the other hand, (10.33) provide the requirements

$$z + \frac{\partial Q^1}{\partial w_1^1} + c\frac{\partial Q^2}{\partial w_1^1} = \frac{\partial Q^1}{\partial w_1^2} + c\left(z + \frac{\partial Q^2}{\partial w_1^2}\right) = \frac{\partial Q^1}{\partial w_1^2} + (c)^2\frac{\partial Q^2}{\partial w_1^1} = 0 \tag{10.41}$$

by using (10.36). If we put

$$z = -\frac{\partial Q^1}{\partial w_1^1} - \frac{\partial Q^2}{\partial w_1^2} = -\left(A^1 + B^1\right)w_1^1 - \left(B^1 + C^2\right)w_1^2, \tag{10.42}$$

then (10.38) is satisfied (a clumsy direct verification).

The above requirements turn to a system of six homogeneous linear equations (not written here) for the six constants A^j, B^j, C^j $(j = 1, 2)$ with determinant $\Delta = c^2(c^2 - 8)$ if the values z, Q^1, Q^2 are inserted and the coefficients of w_1^1 and w_1^2 are compared. The roots $c = 0$ and $c = \pm 2\sqrt{2}$ of the equation $\Delta = 0$ provide rather nontrivial infinitesimal transformation Z, however, we can state only the simplest result for the trivial root $c = 0$ for obvious reason. It reads

$$Q^1 = A^1\left(w_1^1\right)^2, \quad Q^2 = A^2\left(w_1^1\right)^2, \quad z = -A^1 w_1^1, \quad z^1 = 0, \quad z^2 = w_1^1\left(A^2 w_1^1 + A^1 w_1^2\right), \tag{10.43}$$

where A^1, A^2 are arbitrary constants.

Remark 10.3. It follows that investigation of vector fields $Z \in \mathbb{G}$ cannot be regarded for easy task and some new powerful methods are necessary, for example, better use of differential forms (involutive systems) with pseudogroup symmetries of the problem (moving frames).

11. A Few Notes on the Symmetries of Differential Equations

The *external theory* deals with (systems of) differential equations ($\mathfrak{DE}$) that are firmly localized in the jet spaces. This is the common approach and it runs as follows. A given finite system of $\mathfrak{DE}$ is infinitely prolonged in order to ensure the compatibility. In general, this prolongation is a toilsome and delicate task, in particular the "singular solutions" are tacitly passed over. The prolongation procedure is expressed in terms of jet variables and as a result a *fixed subspace* of the (infinite-order) jet space appears which represents the $\mathfrak{DE}$ under consideration. Then the *external symmetries* [2, 3, 6, 7] are such symmetries of the ambient jet space which preserve the subspace. In this sense we may speak of *classical symmetries* (point and contact transformations) and *higher-order symmetries* (which destroy the order of derivatives).

The *internal theory* of $\mathfrak{DE}$ is irrelevant to the jet localization, in particular to the choice of the hierarchy of independent and dependent variables. This point of view is due to E. Cartan and actually the congenial term "diffiety" was introduced in [6, 7]. Alas, these diffieties were defined as objects *locally identical with appropriate external* $\mathfrak{DE}$ restricted to the corresponding subspace of the ambient total jet space. This can hardly be regarded as a coordinate-free (or jet theory-free) approach since the model objects (*external* $\mathfrak{DE}$) and the intertwining mappings (*higher-order symmetries*) essentially need the use of the above hard jet theory mechanisms and concepts.

to Contents, page 1

In reality, the final result of prolongation, the infinitely prolonged $\mathfrak{DE}$, can be alternatively characterized by three simple axioms as follows [8, 9, 24–27].

Let $\mathbf{M}$ be a space *modelled on* $\mathbb{R}^\infty$ (local coordinates $h^1, h^2, \ldots$ as in Sections 1 and 2 above). Denote by $\mathcal{F}(\mathbf{M})$ the *structural module* of all smooth functions f on $\mathbf{M}$ (locally depending on a finite number $m(f)$ of coordinates). Let $\Phi(\mathbf{M})$, $\mathcal{T}(\mathbf{M})$ be the $\mathcal{F}(\mathbf{M})$-modules of all differential 1-forms and vector fields on $\mathbf{M}$, respectively. For every submodule $\Omega \subset \Phi(\mathbf{M})$, we have the "orthogonal" submodule $\Omega^\perp = \mathscr{H} \subset \mathcal{T}(\mathbf{M})$ of all $X \in \mathscr{H}$ such that $\Omega(X) = 0$.

Then an $\mathcal{F}(\mathbf{M})$-submodule $\Omega \subset \Phi(\mathbf{M})$ is called a *diffiety* if the following three requirements are locally satisfied.

($\mathscr{A}$) Ω is of codimension $n < \infty$, equivalent $\mathscr{H}$ is of dimension $n < \infty$.
Here n is the *number of independent variables*. The independent variables provide the complementary module to Ω in $\Phi(\mathbf{M})$ which is not prescribed in advance.

($\mathcal{B}$) $d\Omega \cong 0 \pmod{\Omega}$, equivalent $\mathcal{L}_{\mathscr{H}}\Omega \subset \Omega$, equivalently: $[\mathscr{H}, \mathscr{H}] \subset \mathscr{H}$.
This *Frobenius condition* ensures the classical *passivity requirement*: we deal with the compatible infinite prolongation of differential equations.

($\mathcal{C}$) There exists filtration $\Omega_* : \Omega_0 \subset \Omega_1 \subset \cdots \subset \Omega = \cup\Omega_l$ by finite-dimensional submodules $\Omega_l \subset \Omega$ such that

$$\mathcal{L}_{\mathscr{H}}\Omega_l \subset \Omega_{l+1} \quad (\text{all } l), \quad \Omega_{l+1} = \Omega_l + \mathcal{L}_{\mathscr{H}}\Omega_l \quad (l \text{ large enough}). \tag{11.1}$$

This condition may be expressed in terms of a $\odot\mathscr{H}$-polynomial algebra on the graded module $\oplus\, \Omega_l/\Omega_{l-1}$ (the *Noetherian property*) and ensures the *finite number of dependent variables*. Filtration Ω_* may be capriciously modified. In particular, various localizations of Ω in jet spaces $\Omega(m, n)$ can be easily obtained.

The *internal symmetries* naturally appear. For instance, a vector field $Z \in \mathcal{T}(\mathbf{M})$ is called a *(universal) variation* of diffiety Ω if $\mathcal{L}_Z\Omega \subset \Omega$ and *infinitesimal symmetry* if moreover Z generates a local group, that is, if and only if $Z \in \mathbb{G}$.

Theorem 11.1 (technical theorem). *Let Z be a variation of diffiety Ω. Then $Z \in \mathbb{G}$ if and only if there is a finite-dimensional $\mathcal{F}(M)$-submodule $\Theta \subset \Omega$ such that*

$$\oplus\mathcal{L}^r_{\mathscr{H}}\Theta = \Omega, \qquad \dim \oplus\mathcal{L}^r_Z\Theta < \infty. \tag{11.2}$$

This is exactly counterpart to Theorem 5.6: submodule $\Theta \subset \Omega$ stands here for the previous submodule $\Omega_0 \subset \Omega(m, n)$. We postpone the proof of Theorem 11.1 together with applications to some convenient occasion.

Remark 11.2. There may exist conical symmetries Z of a diffiety Ω, however, they are all lying in $\mathscr{H}$ and generate just the *Cauchy characteristics* of the diffiety [9, page 155].

We conclude with two examples of internal theory of underdetermined ordinary differential equations. The reasonings to follow can be carried over quite general diffieties without any change.

Example 11.3. Let us deal with the *Monge equation*

$$\frac{dx}{dt} = f\left(t, x, y, \frac{dy}{dt}\right). \tag{11.3}$$

The prolongation can be represented as the Pfaffian system

$$\mathrm{d}x - f(t, x, y, y')\mathrm{d}t = 0, \quad \mathrm{d}y - y'\mathrm{d}t = 0, \quad \mathrm{d}y' - y''\mathrm{d}t = 0, \ldots. \tag{11.4}$$

Within the framework of diffieties, we introduce space $\mathbf{M}$ with coordinates

$$t, x_0, y_0, y_1, y_2, \ldots \tag{11.5}$$

and submodule $\Omega \subset \Phi(\mathbf{M})$ with generators

$$\mathrm{d}x_0 - f\mathrm{d}t, \quad (\omega_r =)\mathrm{d}y_r - y_{r+1}\mathrm{d}t \quad (r = 0, 1, \ldots;\ f = f(t, x_0, y_0, y_1)). \tag{11.6}$$

Clearly $\mathcal{H} = \Omega^\perp \subset \mathcal{T}(\mathbf{M})$ is one-dimensional subspace including the vector field

$$D = \frac{\partial}{\partial t} + f\frac{\partial}{\partial x_0} + \sum y_{r+1}\frac{\partial}{\partial y_r}. \tag{11.7}$$

One can easily find that we have a diffiety. ($\mathcal{A}$ and $\mathcal{B}$ are trivially satisfied. The *common order preserving filtrations* where Ω_l involves $\mathrm{d}x_0 - f\mathrm{d}t$ and ω_r with $r \le l$ is enough for $\mathcal{C}$.)

We introduce a new (*standard* [9]) *filtration* $\overline{\Omega}_*$ where the submodule $\overline{\Omega}_l \subset \Omega$ is generated by the forms

$$\vartheta_0 = \mathrm{d}x_0 - f\mathrm{d}t - \frac{\partial f}{\partial y_1}\omega_0, \omega_r \quad (r \le l - 1). \tag{11.8}$$

This is indeed a filtration since

$$\begin{aligned}\mathcal{L}_D\vartheta_0 &= \mathrm{d}f - Df\mathrm{d}t - D\frac{\partial f}{\partial y_1}\cdot\omega_0 - \frac{\partial f}{\partial y_1}\omega_1 = \frac{\partial f}{\partial x_0}(\mathrm{d}x_0 - f\mathrm{d}t) + \left(\frac{\partial f}{\partial y_0} - D\frac{\partial f}{\partial y_1}\right)\omega_0 \\ &= \frac{\partial f}{\partial x_0}\vartheta_0 + A\omega_0 \quad \left(A = \frac{\partial f}{\partial y_0} + \frac{\partial f}{\partial x_0}\frac{\partial f}{\partial y_1} - D\frac{\partial f}{\partial y_1}\right)\end{aligned} \tag{11.9}$$

and (trivially) $\mathcal{L}_D\omega_r = \omega_{r+1}$. Assuming $A \neq 0$ from now on (this is satisfied if $f_{y_1 y_1} \neq 0$) *every module $\overline{\Omega}_l$ is generated by the forms $\vartheta_r = \mathcal{L}_D^r\vartheta_0$ $(r \le l)$.*

The forms ϑ_r satisfy the recurrence $\mathcal{L}_D\vartheta_r = \vartheta_{r+1}$. Then the formula

$$\vartheta_{r+1} = \mathcal{L}_D\vartheta_r = D\rfloor\mathrm{d}\vartheta_r + \mathrm{d}\vartheta_r(D) = D\rfloor\mathrm{d}\vartheta_r \tag{11.10}$$

to Contents, page 1

implies the congruence $\mathrm{d}\vartheta_r \cong \mathrm{d}t \wedge \vartheta_{r+1}$ (mod $\Omega \wedge \Omega$). Let

$$Z = z\frac{\partial}{\partial t} + z^0\frac{\partial}{\partial x_0} + \sum z_r\frac{\partial}{\partial y_r} \tag{11.11}$$

be a variation of Ω in the common sense $\mathcal{L}_Z\Omega \subset \Omega$. This inclusion is equivalent to the congruence

$$\mathcal{L}_Z\vartheta_r = Z\rfloor \mathrm{d}\vartheta_r + \mathrm{d}\vartheta_r(Z) \cong -\vartheta_{r+1}(Z)\mathrm{d}t + D\vartheta_r(Z)\mathrm{d}t = 0 \,(\mathrm{mod}\ \Omega) \tag{11.12}$$

whence to the recurrence

$$\vartheta_{r+1}(Z) = D\vartheta_r(Z) \tag{11.13}$$

quite analogous to the recurrence (5.10), see Remark 5.3. It follows that the functions

$$z = Zt = \mathrm{d}t(Z), \qquad g = \vartheta_0(Z) \tag{11.14}$$

can be quite arbitrarily chosen. Then functions $\vartheta_r(Z) = D^r g$ are determined and we obtain *quite explicit formulae for the variation* Z. In more detail

$$\begin{gathered} g = \vartheta_0(Z) = \left(\mathrm{d}x_0 - f\mathrm{d}t - \frac{\partial f}{\partial y_1}\omega_0\right)(Z) = z^0 - fz - \frac{\partial f}{\partial y_1}(z_0 - y_1 z), \\ Dg = \vartheta_1(Z)\left(\frac{\partial f}{\partial x_0}\vartheta_0 + A\omega_0\right)(Z) = \frac{\partial f}{\partial x_0}g + A(z_0 - y_1 z) \end{gathered} \tag{11.15}$$

and these equations determine coefficients z^0 and z_0 in terms of functions z and g. Coefficients z_r $(r \geq 1)$ follow by prolongation (not stated here). If moreover

$$\dim \{\mathcal{L}^r_Z\vartheta_0\}_{r\in\mathbb{N}} < \infty \tag{11.16}$$

we have infinitesimal symmetry $Z \in \mathbb{G}$, see Theorem 11.1.

Example 11.4. Let us deal with the *Hilbert-Cartan equation* [3]

$$\frac{dy}{dt} = \left(\frac{d^2x}{dt^2}\right)^2. \tag{11.17}$$

Passing to the diffiety, we introduce space **M** with coordinates

$$t, x_0, x_1, y_0, y_1, y_2, \ldots \tag{11.18}$$

to Contents, page 1

and submodule $\Omega \subset \Phi(\mathbf{M})$ generated by forms

$$\mathrm{d}x_0 - x_1\mathrm{d}t, \quad \mathrm{d}x_1 - \sqrt{y_1}\mathrm{d}t, \quad (\omega_r =)\mathrm{d}y_r - y_{r+1}\mathrm{d}t \quad (r = 0, 1, \ldots). \tag{11.19}$$

The submodule $\mathscr{H} = \Omega^{\perp} \subset \mathfrak{T}(\mathbf{M})$ is generated by the vector field

$$D = \frac{\partial}{\partial t} + x_1\frac{\partial}{\partial x_0} + \sqrt{y_1}\frac{\partial}{\partial x_1} + \sum y_{r+1}\frac{\partial}{\partial y_r}. \tag{11.20}$$

We introduce the form

$$\vartheta_0 = \mathrm{d}x_0 - x_1\mathrm{d}t + B\left\{\mathrm{d}x_1 - \sqrt{y_1}\mathrm{d}t - \frac{1}{2\sqrt{y_1}}\omega_0\right\} \quad \left(B = \frac{1/\sqrt{y_1}}{D(1/\sqrt{y_1})}\right) \tag{11.21}$$

and moreover the forms

$$\begin{aligned}
&\vartheta_1 = \mathcal{L}_D\vartheta_0 = (1 + DB)\{\cdots\}, \\
&\vartheta_2 = \mathcal{L}_D\vartheta_1 = D^2B\{\cdots\} - C\omega_0 \quad \left(C = (1 + DB)D\frac{1}{2\sqrt{y_1}}\right), \\
&\vartheta_3 = \cdots + C\omega_1, \\
&\vartheta_4 = \cdots + C\omega_2, \\
&\quad \vdots
\end{aligned} \tag{11.22}$$

Assuming $C \neq 0$, we have a standard filtration $\overline{\Omega}_*$ where the submodules $\overline{\Omega}_l \subset \Omega$ are generated by forms ϑ_r $(r \le l)$. Explicit formulae for variations

$$Z = z\frac{\partial}{\partial t} + z^0\frac{\partial}{\partial x_0} + z^1\frac{\partial}{\partial x_1} + \sum z_r\frac{\partial}{\partial y_r} \tag{11.23}$$

can be obtained analogously as in Example 11.3 (and are omitted here). Functions z and $g = \vartheta_0(Z)$ can be arbitrarily chosen. Condition (11.16) ensures $Z \in \mathbb{G}$.

Appendix

For the convenience of reader, we survey some results [9, 18, 19] on the modules Adj. Our reasonings are carried out in the space $\mathbb{R}^n$ and will be true *locally* near *generic points*.

Let Θ be a given module of 1-forms and $A(\Theta)$ the module of all vector fields X such that $\mathcal{L}_{fX}\Theta \subset \Theta$ for all functions f, see [9]. Clearly

$$\mathcal{L}_{[X,Z]}\Theta = (\mathcal{L}_X\mathcal{L}_Z - \mathcal{L}_Z\mathcal{L}_X)\Theta \subset \Theta \quad (X, Z \in A(\Theta)) \tag{A.1}$$

to Contents, page 1

and it follows that identity

$$f[X,Y] = [X,Z] + Xf \cdot Y \quad \big(X,Y \in A(\Theta);\ Z = fY\big) \tag{A.2}$$

implies $\mathcal{L}_{f[X,Y]}\Theta \subset \Theta$ whence $[A(\Theta), A(\Theta)] \subset A(\Theta)$.

Let Θ be of a finite dimension I. The Frobenius theorem can be applied, and it follows that module $\operatorname{Adj}\Theta = A(\Theta)^{\perp}$ (of all forms φ satisfying $\varphi(A(\Theta)) = 0$) has a certain basis $\mathrm{d}f^1,\dots,\mathrm{d}f^K$ $(K \geq I)$.

On the other hand, identity

$$\mathcal{L}_{fX}\vartheta = \quad fX\rfloor \mathrm{d}\vartheta + \mathrm{d}\big(f\vartheta(X)\big) = f\mathcal{L}_X\vartheta + \vartheta(X)\vartheta \tag{A.3}$$

implies that $X \in A(\Theta)$ if and only if

$$\vartheta(X) = 0,\ X\rfloor \mathrm{d}\vartheta \in \Theta \quad (\vartheta \in \Theta) \tag{A.4}$$

which is the classical definition, see [2]. In particular $\Theta \subset \operatorname{Adj}\Theta$ so we may suppose the generators

$$\vartheta^i = \mathrm{d}f^i + g^i_{I+1}\mathrm{d}f^{I+1} + \cdots + g^i_K\mathrm{d}f^K \in \Theta \quad (i = 1,\dots,I) \tag{A.5}$$

of module Θ. Recall that $Xf^k = 0 \ \ (k = 1,\dots,K;\ X \in A(\Theta))$ whence

$$\mathcal{L}_X\vartheta^i = Xg^i_{I+1}\mathrm{d}f^{I+1} + \cdots + Xg^i_K\mathrm{d}f^K \in \Theta \tag{A.6}$$

and this implies $Xg^i_{I+1} = \cdots = Xg^i_K = 0$. It follows that

$$\mathrm{d}g^i_{I+1},\dots,\mathrm{d}g^i_K \in \operatorname{Adj}\Theta \quad (i = 1,\dots,I) \tag{A.7}$$

and therefore all coefficients g^i_k depend only on variables $f^1,\dots,f^K$.

Acknowledgment

This research has been conducted at the Department of Mathematics as part of the research project CEZ: Progressive reliable and durable structures, MSM 0021630519.

References

[1] R. L. Anderson and N. H. Ibragimov, *Lie-Bäcklund Transformations in Applications*, vol. 1 of *SIAM Studies in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 1979.

[2] I. M. Anderson, N. Kamran, and P. J. Olver, "Internal, external, and generalized symmetries," *Advances in Mathematics*, vol. 100, no. 1, pp. 53–100, 1993.

[3] N. Kamran, *Selected Topics in the Geometrical Study of Differential Equations*, vol. 96 of *CBMS Regional Conference Series in Mathematics*, American Mathematical Society, Washington, DC, USA, 2002.

to Contents, page 1

[4] P. J. Olver, *Applications of Lie Groups to Differential Equations*, vol. 107 of *Graduate Texts in Mathematics*, Springer, New York, NY, USA, 1986.
[5] P. J. Olver, *Equivalence, Invariants, and Symmetry*, Cambridge University Press, Cambridge, UK, 1995.
[6] A. M. Vinogradov, I. S. Krasil'shchik, and V. V. Lychagin, *Introduction to the Theory of Nonlinear Differential Equations*, "Nauka", Moscow, Russia, 1986.
[7] A. M. Vinogradov, *Cohomological Analysis of Partial Differential Equations and Secondary Calculus*, vol. 204 of *Translations of Mathematical Monographs*, American Mathematical Society, Providence, RI, USA, 2001.
[8] J. Chrastina, "From elementary algebra to Bäcklund transformations," *Czechoslovak Mathematical Journal*, vol. 40(115), no. 2, pp. 239–257, 1990.
[9] J. Chrastina, *The Formal Theory of Differential Equations*, vol. 6 of *Folia Facultatis Scientiarium Naturalium Universitatis Masarykianae Brunensis. Mathematica*, Masaryk University, Brno, Czech Republic, 1998.
[10] V. Tryhuk and V. Chrastinová, "Automorphisms of curves," *Journal of Nonlinear Mathematical Physics*, vol. 16, no. 3, pp. 259–281, 2009.
[11] V. Chrastinová and V. Tryhuk, "Generalized contact transformations," *Journal of Applied Mathematics, Statistics and Informatics*, vol. 3, no. 1, pp. 47–62, 2007.
[12] V. Tryhuk and V. Chrastinová, "On the mapping of jet spaces," *Journal of Nonlinear Mathematical Physics*, vol. 17, no. 3, pp. 293–310, 2010.
[13] V. Chrastinová and V. Tryhuk, "Automorphisms of submanifolds," *Advances in Difference Equations*, vol. 2010, Article ID 202731, 26 pages, 2010.
[14] N. H. Ibragimov, *Transformation Groups Applied to Mathematical Physics*, Mathematics and Its Applications, D. Reidel, Dordrecht, The Netherlands, 1985.
[15] J. Mikeš, A. Vanžurová, and I. Hinterleitner, *Geodesic Mappings and Some Generalizations*, Palacký University Olomouc, Faculty of Science, Olomouc, Czech Republic, 2009.
[16] V. N. Chetverikov, "On the structure of integrable $\mathcal{C}$-fields," *Differential Geometry and Its Applications*, vol. 1, no. 4, pp. 309–325, 1991.
[17] L. P. Eisenhart, *Continuous Groups of Transformations*, Princeton University Press, 1933.
[18] É. Cartan, *Les systèmes extérieurs et leurs applications géométriques*, Hermann, Paris, France, 1946.
[19] É. Cartan, *Les systèmes extérieurs et leurs applications géométriques*, Actualits scientifiques et industrielles. No. 994, Hermann, Paris, France, 2nd edition, 1971.
[20] V. Guillemin and S. Sternberg, *Geometric Asymptotics*, Mathematical Surveys, No. 1, American Mathematical Society, Providence, RI, USA, 1977.
[21] A. Kushner, V. Lychagin, and V. Rubtsov, *Contact Geometry and Non-Linear Differential Equations*, vol. 101 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge, UK, 2007.
[22] É. Goursat, *Leçons sur l'intégration des équations aux dérivées partielles premier ordre*, Zweite Auflage, Paris, France, 1891.
[23] É. Goursat, *Leçons sur l'intégration des équations aux dérivées partielles premier ordre*, J. Hermann, Paris, France, 1920.
[24] J. Chrastina, "What the differential equations should be," in *Proceedings of the Conference on Differential Geometry and Its Applications, Part 2*, pp. 41–50, Univ. J. E. Purkyně, Brno, Czech Republic, 1984.
[25] J. Chrastina, "On formal theory of differential equations. I," *Časopis Pro Pěstování Matematiky*, vol. 111, no. 4, pp. 353–383, 1986.
[26] J. Chrastina, "On formal theory of differential equations. II," *Časopis Pro Pěstování Matematiky*, vol. 114, no. 1, pp. 60–105, 1989.
[27] J. Chrastina, "On formal theory of differential equations. III," *Mathematica Bohemica*, vol. 116, no. 1, pp. 60–90, 1991.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 285040, 15 pages
doi:10.1155/2011/285040

Research Article

The Local Strong and Weak Solutions for a Nonlinear Dissipative Camassa-Holm Equation

Shaoyong Lai

Department of Applied Mathematics, Southwestern University of Finance and Economics, Chengdu, 610074, China

Correspondence should be addressed to Shaoyong Lai, laishaoy@swufe.edu.cn

Received 17 January 2011; Revised 3 August 2011; Accepted 8 August 2011

Academic Editor: Yuri V. Rogovchenko

Using the Kato theorem for abstract differential equations, the local well-posedness of the solution for a nonlinear dissipative Camassa-Holm equation is established in space $C([0,T), H^s(R)) \cap C^1([0,T), H^{s-1}(R))$ with $s > 3/2$. In addition, a sufficient condition for the existence of weak solutions of the equation in lower order Sobolev space $H^s(R)$ with $1 \le s \le 3/2$ is developed.

1. Introduction

Camassa and Holm [1] used the Hamiltonian method to derive a completely integrable wave equation

$$u_t - u_{xxt} + 2ku_x + 3uu_x = 2u_x u_{xx} + uu_{xxx}, \tag{1.1}$$

by retaining two terms that are usually neglected in the small amplitude, shallow water limit. Its alternative derivation as a model for water waves can be found in Constantin and Lannes [2] and Johnson [3]. Equation (1.1) also models wave current interaction [4], while Dai [5] derived it as a model in elasticity (see Constantin and Strauss [6]). Moreover, it was pointed out in Lakshmanan [7] that the Camassa-Holm equation (1.1) could be relevant to the modeling of tsunami waves (see Constantin and Johnson [8]).

In fact, a huge amount of work has been carried out to investigate the dynamic properties of (1.1). For $k = 0$, (1.1) has traveling wave solutions of the form $c\,e^{-|x-ct|}$, called peakons, which capture the main feature of the exact traveling wave solutions of greatest height of the governing equations (see [9–11]). For $k > 0$, its solitary waves are stable solitons [6, 11]. It was shown in [12–14] that the inverse spectral or scattering approach was

to Contents, page 1

a powerful tool to handle Camassa-Holm equation. Equation (1.1) is a completely integrable infinite-dimensional Hamiltonian system (in the sense that for a large class of initial data, the flow is equivalent to a linear flow at constant speed [15]). It should be emphasized that (1.1) gives rise to geodesic flow of a certain invariant metric on the Bott-Virasoro group (see [16, 17]), and this geometric illustration leads to a proof that the Least Action Principle holds. It is worthwhile to mention that Xin and Zhang [18] proved that the global existence of the weak solution in the energy space $H^1(R)$ without any sign conditions on the initial value, and the uniqueness of this weak solution is obtained under some conditions on the solution [19]. Coclite et al. [20] extended the analysis presented in [18, 19] and obtained many useful dynamic properties to other equations (also see [21–24]). Li and Olver [25] established the local well-posedness in the Sobolev space $H^s(R)$ with $s > 3/2$ for (1.1) and gave conditions on the initial data that lead to finite time blowup of certain solutions. It was shown in Constantin and Escher [26] that the blowup occurs in the form of breaking waves, namely, the solution remains bounded but its slope becomes unbounded in finite time. After wave breaking, the solution can be continued uniquely either as a global conservative weak solution [21] or a global dissipative solution [22]. For peakons, these possibilities are explicitly illustrated in the paper [27]. For other methods to handle the problems relating to various dynamic properties of the Camassa-Holm equation and other shallow water models, the reader is referred to [10, 28–32] and the references therein.

In this paper, motivated by the work in [25, 33], we study the following generalized Camassa-Holm equation

$$u_t - u_{txx} + 2ku_x + au^m u_x = 2u_x u_{xx} + uu_{xxx} + \beta\partial_x\left[(u_x)^N\right], \tag{1.2}$$

where $m \geq 1$ and $N \geq 1$ are natural numbers, and a, k, and β are arbitrary constants. Obviously, (1.2) reduces to (1.1) if we set $a = 3$, $m = 1$, and $\beta = 0$. Actually, Wu and Yin [34] consider a nonlinearly dissipative Camassa-Holm equation which includes a nonlinearly dissipative term $L(u)$, where L is a differential operator or a quasidifferential operator. Therefore, we can regard the term $\beta\partial_x[(u_x)^N]$ as a nonlinearly dissipative term for the dissipative Camassa-Holm equation (1.2).

Due to the term $\beta\partial_x[(u_x)^N]$ in (1.2), the conservation laws in previous works [10, 25] for (1.1) lose their powers to obtain some bounded estimates of the solution for (1.2). A new conservation law different from those presented in [10, 25] will be established to prove the local existence and uniqueness of the solution to (2.3) subject to initial value $u_0(x) \in H^s(R)$ with $s > 3/2$. We should address that all the generalized versions of the Camassa-Holm equation in previous works (see [17, 25, 34]) do not involve the nonlinear term $\partial_x[(u_x)^N]$. Lai and Wu [33] only studied a generalized Camassa-Holm equation in the case where $\beta \geq 0$ and N is an odd number. Namely, (1.2) with $\beta < 0$ and arbitrary positive integer N was not investigated in [33].

The main tasks of this paper are two-fold. Firstly, by using the Kato theorem for abstract differential equations, we establish the local existence and uniqueness of solutions for (1.2) with any β and arbitrary positive integer N in space $C([0,T), H^s(R)) \bigcap C^1([0,T), H^{s-1}(R))$ with $s > 3/2$. Secondly, it is shown that the existence of weak solutions in lower order Sobolev space $H^s(R)$ with $1 \leq s \leq 3/2$. The ideas of proving the second result come from those presented in Li and Olver [25].

to Contents, page 1

2. Main Results

Firstly, we give some notation.

The space of all infinitely differentiable functions $\phi(t,x)$ with compact support in $[0,+\infty)\times R$ is denoted by C_0^∞. $L^p = L^p(R)$ $(1 \le p < +\infty)$ is the space of all measurable functions h such that $\|h\|_{L^p}^p = \int_R |h(t,x)|^p dx < \infty$. We define $L^\infty = L^\infty(R)$ with the standard norm $\|h\|_{L^\infty} = \inf_{m(e)=0}\sup_{x\in R\setminus e}|h(t,x)|$. For any real number s, $H^s = H^s(R)$ denotes the Sobolev space with the norm defined by

$$\|h\|_{H^s} = \left(\int_R \left(1+|\xi|^2\right)^s \left|\widehat{h}(t,\xi)\right|^2 d\xi\right)^{1/2} < \infty, \tag{2.1}$$

where $\widehat{h}(t,\xi) = \int_R e^{-ix\xi}h(t,x)dx$.

For $T > 0$ and nonnegative number s, $C([0,T);H^s(R))$ denotes the Frechet space of all continuous H^s-valued functions on $[0,T)$. We set $\Lambda = (1-\partial_x^2)^{1/2}$.

In order to study the existence of solutions for (1.2), we consider its Cauchy problem in the form

$$\begin{aligned} u_t - u_{txx} &= -2ku_x - \frac{a}{m+1}\left(u^{m+1}\right)_x + 2u_xu_{xx} + uu_{xxx} + \beta\partial_x\left[(u_x)^N\right] \\ &= -ku_x - \frac{a}{m+1}\left(u^{m+1}\right)_x + \frac{1}{2}\partial_x^3u^2 - \frac{1}{2}\partial_x\left(u_x^2\right) + \beta\partial_x\left[(u_x)^N\right], \\ u(0,x) &= u_0(x), \end{aligned} \tag{2.2}$$

which is equivalent to

$$\begin{gathered} u_t + uu_x = \Lambda^{-2}\left[-ku - \frac{a}{m+1}\left(u^{m+1}\right)\right]_x + \Lambda^{-2}(uu_x) - \frac{1}{2}\Lambda^{-2}\partial_x\left(u_x^2\right) + \beta\Lambda^{-2}\partial_x\left[(u_x)^N\right], \\ u(0,x) = u_0(x). \end{gathered} \tag{2.3}$$

Now, we state our main results.

Theorem 2.1. *Let $u_0(x) \in H^s(R)$ with $s > 3/2$. Then problem* (2.2) *or problem* (2.3) *has a unique solution $u(t,x) \in C([0,T);H^s(R))\bigcap C^1([0,T);H^{s-1}(R))$ where $T > 0$ depends on $\|u_0\|_{H^s(R)}$.*

Theorem 2.2. *Suppose that $u_0(x) \in H^s$ with $1 \le s \le 3/2$ and $\|u_{0x}\|_{L^\infty} < \infty$. Then there exists a $T > 0$ such that* (1.2) *subject to initial value $u_0(x)$ has a weak solution $u(t,x) \in L^2([0,T],H^s)$ in the sense of distribution and $u_x \in L^\infty([0,T]\times R)$.*

3. Local Well-Posedness

We consider the abstract quasilinear evolution equation

$$\frac{dv}{dt} + A(v)v = f(v), \quad t \ge 0,\ v(0) = v_0. \tag{3.1}$$

to Contents, page 1

Let X and Y be Hilbert spaces such that Y is continuously and densely embedded in X, and let $Q : Y \to X$ be a topological isomorphism. Let $L(Y, X)$ be the space of all bounded linear operators from Y to X. If $X = Y$, we denote this space by $L(X)$. We state the following conditions in which $\rho_1,\ \rho_2,\ \rho_3$, and ρ_4 are constants depending on $\max\{\|y\|_Y, \|z\|_Y\}$.

(i) $A(y) \in L(Y, X)$ for $y \in X$ with

$$\left\|(A(y) - A(z))w\right\|_X \leq \rho_1 \left\|y - z\right\|_X \|w\|_Y, \quad y, z, w \in Y, \tag{3.2}$$

and $A(y) \in G(X, 1, \beta)$ (i.e., $A(y)$ is quasi-m-accretive), uniformly on bounded sets in Y.

(ii) $QA(y)Q^{-1} = A(y) + B(y)$, where $B(y) \in L(X)$ is bounded, uniformly on bounded sets in Y. Moreover,

$$\left\|(B(y) - B(z))w\right\|_X \leq \rho_2 \left\|y - z\right\|_Y \|w\|_X, \quad y, z \in Y,\ w \in X. \tag{3.3}$$

(iii) $f : Y \to Y$ extends to a map from X into X is bounded on bounded sets in Y, and satisfies

$$\begin{aligned} \left\|f(y) - f(z)\right\|_Y &\leq \rho_3 \left\|y - z\right\|_Y, \quad y, z \in Y, \\ \left\|f(y) - f(z)\right\|_X &\leq \rho_4 \left\|y - z\right\|_X, \quad y, z \in Y. \end{aligned} \tag{3.4}$$

Kato Theorem (see [35])

Assume that (i), (ii), and (iii) hold. If $v_0 \in Y$, there is a maximal $T > 0$ depending only on $\|v_0\|_Y$, and a unique solution v to problem (3.1) such that

$$v = v(\cdot, v_0) \in C([0, T); Y) \bigcap C^1([0, T); X). \tag{3.5}$$

Moreover, the map $v_0 \to v(\cdot, v_0)$ is a continuous map from Y to the space

$$C([0, T); Y) \bigcap C^1([0, T); X). \tag{3.6}$$

For problem (2.3), we set $A(u) = u\partial_x$, $Y = H^s(R)$, $X = H^{s-1}(R)$, $\Lambda = (1 - \partial_x^2)^{1/2}$,

$$f(u) = \Lambda^{-2}\left[-ku - \frac{a}{m+1}\left(u^{m+1}\right)\right]_x + \Lambda^{-2}(uu_x) - \frac{1}{2}\Lambda^{-2}\partial_x\left(u_x^2\right) + \beta\Lambda^{-2}\partial_x\left[(u_x)^N\right], \tag{3.7}$$

and $Q = \Lambda$. In order to prove Theorem 2.1, we only need to check that $A(u)$ and $f(u)$ satisfy assumptions (i)–(iii).

Lemma 3.1. *The operator $A(u) = u\partial_x$ with $u \in H^s(R)$, $s > 3/2$ belongs to $G(H^{s-1}, 1, \beta)$.*

to Contents, page 1

Lemma 3.2. *Let $A(u) = u\partial_x$ with $u \in H^s$ and $s > 3/2$. Then $A(u) \in L(H^s, H^{s-1})$ for all $u \in H^s$. Moreover,*

$$\|(A(u) - A(z))w\|_{H^{s-1}} \le \rho_1 \|u - z\|_{H^{s-1}} \|w\|_{H^s}, \quad u, z, w \in H^s(R). \tag{3.8}$$

Lemma 3.3. *For $s > 3/2$, $u, z \in H^s$ and $w \in H^{s-1}$, it holds that $B(u) = [\Lambda, u\partial_x]\Lambda^{-1} \in L(H^{s-1})$ for $u \in H^s$ and*

$$\|(B(u) - B(z))w\|_{H^{s-1}} \le \rho_2 \|u - z\|_{H^s} \|w\|_{H^{s-1}}. \tag{3.9}$$

Proofs of the above Lemmas 3.1–3.3 can be found in [29] or [31].

Lemma 3.4 (see [35]). *Let r and q be real numbers such that $-r < q \le r$. Then*

$$\begin{aligned} \|uv\|_{H^q} &\le c\|u\|_{H^r}\|v\|_{H^q}, \quad \text{if } r > \frac{1}{2}, \\ \|uv\|_{H^{r+q-1/2}} &\le c\|u\|_{H^r}\|v\|_{H^q}, \quad \text{if } r < \frac{1}{2}. \end{aligned} \tag{3.10}$$

Lemma 3.5. *Let $u, z \in H^s$ with $s > 3/2$, then $f(u)$ is bounded on bounded sets in H^s and satisfies*

$$\|f(u) - f(z)\|_{H^s} \le \rho_3 \|u - z\|_{H^s}, \tag{3.11}$$

$$\|f(u) - f(z)\|_{H^{s-1}} \le \rho_4 \|u - z\|_{H^{s-1}}. \tag{3.12}$$

Proof. Using the algebra property of the space H^{s_0} with $s_0 > 1/2$, we have

$$\begin{aligned} &\|f(u) - f(z)\|_{H^s} \\ &\le c\Bigg[\left\|\Lambda^{-2}\left(\left[-ku - \frac{a}{m+1}\left(u^{m+1}\right)\right]_x - \left[-kz - \frac{a}{m+1}\left(z^{m+1}\right)\right]_x\right)\right\|_{H^s} \\ &\qquad + \left\|\Lambda^{-2}(uu_x - zz_x)\right\|_{H^s} + \left\|\Lambda^{-2}\partial_x\left(u_x^2 - z_x^2\right)\right\|_{H^s} + \left\|\Lambda^{-2}\partial_x\left[(u_x)^N\right] - \Lambda^{-2}\partial_x\left[(z_x)^N\right]\right\|_{H^s}\Bigg] \\ &\le c\Bigg[\|u - z\|_{H^{s-1}} + \left\|u^{m+1} - z^{m+1}\right\|_{H^{s-1}} + \|uu_x - zz_x\|_{H^{s-1}} + \left\|u_x^2 - z_x^2\right\|_{H^{s-1}} \\ &\qquad + \left\|(u_x)^N - (z_x)^N\right\|_{H^{s-1}}\Bigg] \\ &\le c\|u - z\|_{H^s}\left[1 + \sum_{j=0}^{m} \|u\|_{H^s}^{m-j}\|z\|_{H^s}^{j} + \|u\|_{H^s} + \|z\|_{H^s} + \sum_{j=0}^{N-1} \|u_x\|_{H^{s-1}}^{N-j}\|z_x\|_{H^{s-1}}^{j}\right] \\ &\le \rho_3 \|u - z\|_{H^s}, \end{aligned} \tag{3.13}$$

from which we obtain (3.11).

to Contents, page 1

Applying Lemma 3.4, $uu_x = (1/2)(u^2)_x$, $s > 3/2$, $\|u\|_{L^\infty} \le c\|u\|_{H^{s-1}}$ and $\|u_x\|_{L^\infty} \le c\|u\|_{H^s}$, we get

$$\begin{aligned}
&\|f(u) - f(z)\|_{H^{s-1}} \\
&\quad \le c\left[\|u - z\|_{H^{s-2}} + \left\|u^{m+1} - z^{m+1}\right\|_{H^{s-2}} + \left\|u^2 - z^2\right\|_{H^{s-2}}\right. \\
&\qquad \left. + \|(u_x - z_x)(u_x + z_x)\|_{H^{s-2}} + \left\|(u_x - z_x)\sum_{j=0}^{N-1} u_x^{N-1-j} z_x^j\right\|_{H^{s-2}}\right] \\
&\quad \le c\|u - z\|_{H^{s-1}}\left[1 + \sum_{j=0}^{m} \|u\|_{H^{s-1}}^{m-j}\|z\|_{H^{s-1}}^{j} + \|u\|_{H^{s-1}} + \|z\|_{H^{s-1}}\right. \\
&\qquad \left. + \|u\|_{H^s} + \|z\|_{H^s} + \sum_{j=0}^{N-1} \|u_x\|_{H^{s-1}}^{N-j}\|z_x\|_{H^{s-1}}^{j}\right] \\
&\quad \le \rho_4\|u - z\|_{H^{s-1}},
\end{aligned} \tag{3.14}$$

which completes the proof of (3.12). □

Proof of Theorem 2.1. Using the Kato Theorem, Lemmas 3.1–3.3, and 3.5, we know that system (2.2) or problem (2.3) has a unique solution

$$u(t,x) \in C([0,T); H^s(R)) \bigcap C^1\left([0,T); H^{s-1}(R)\right). \tag{3.15}$$

□

4. Existence of Weak Solutions

For $s \ge 2$, using the first equation of system (2.2) derives

$$\frac{d}{dt}\int_R \left(u^2 + u_x^2 + 2\beta\int_0^t u_x^{N+1} d\tau\right) dx = 0, \tag{4.1}$$

from which we have the conservation law

$$\int_R \left(u^2 + u_x^2 + 2\beta\int_0^t u_x^{N+1} d\tau\right) dx = \int_R \left(u_0^2 + u_{0x}^2\right) dx. \tag{4.2}$$

Lemma 4.1 (Kato and Ponce [36]). *If $r > 0$, then $H^r \bigcap L^\infty$ is an algebra. Moreover,*

$$\|uv\|_r \le c(\|u\|_{L^\infty}\|v\|_r + \|u\|_r\|v\|_{L^\infty}), \tag{4.3}$$

where c is a constant depending only on r.

to Contents, page 1

Lemma 4.2 (Kato and Ponce [36]). *Let $r > 0$. If $u \in H^r \bigcap W^{1,\infty}$ and $v \in H^{r-1} \bigcap L^\infty$, then*

$$\|[\Lambda^r, u]v\|_{L^2} \leq c\left(\|\partial_x u\|_{L^\infty} \left\|\Lambda^{r-1} v\right\|_{L^2} + \|\Lambda^r u\|_{L^2} \|v\|_{L^\infty}\right). \tag{4.4}$$

Lemma 4.3. *Let $s \geq 2$ and the function $u(t, x)$ is a solution of problem (2.2) and the initial data $u_0(x) \in H^s(R)$. Then the following inequality holds*

$$\|u\|_{L^\infty} \leq \|u\|_{H^1} \leq \|u_0\|_{H^1} e^{|\beta| \int_0^t \|u_x\|_{L^\infty}^{N-1} d\tau}. \tag{4.5}$$

For $q \in (0, s-1]$, there is a constant c, which only depends on m, N, k, a, and β, such that

$$\begin{aligned} \int_R \left(\Lambda^{q+1} u\right)^2 dx \leq& \int_R \left(\Lambda^{q+1} u_0\right)^2 dx + c\int_0^t \|u_x\|_{L^\infty} \|u\|_{H^{q+1}}^2 \left(1 + \|u\|_{L^\infty}^{m-1}\right) d\tau \\ &+ c\int_0^t \|u\|_{H^{q+1}}^2 \|u_x\|_{L^\infty}^{N-1} d\tau. \end{aligned} \tag{4.6}$$

For $q \in [0, s-1]$, there is a constant c, which only depends on m, N, k, a, and β, such that

$$\|u_t\|_{H^q} \leq c\|u\|_{H^{q+1}} \left(1 + \left(1 + \|u\|_{L^\infty}^{m-1}\right)\|u\|_{H^1} + \|u_x\|_{L^\infty}^{N-1}\right). \tag{4.7}$$

Proof. Using $\|u\|_{H^1}^2 = \int_R (u^2 + u_x^2) dx$ and (4.2) derives (4.5).

Using $\partial_x^2 = -\Lambda^2 + 1$ and the Parseval equality gives rise to

$$\int_R \Lambda^q u \Lambda^q \partial_x^3 \left(u^2\right) dx = -2\int_R \left(\Lambda^{q+1} u\right) \Lambda^{q+1}(u u_x) dx + 2\int_R (\Lambda^q u) \Lambda^q (u u_x) dx. \tag{4.8}$$

For $q \in (0, s-1]$, applying $(\Lambda^q u)\Lambda^q$ to both sides of the first equation of system (2.3) and integrating with respect to x by parts, we have the identity

$$\begin{aligned} \frac{1}{2}\frac{d}{dt} \int_R \left((\Lambda^q u)^2 + (\Lambda^q u_x)^2\right) dx =& -a\int_R (\Lambda^q u) \Lambda^q (u^m u_x) dx \\ &- \int_R \left(\Lambda^{q+1} u\right) \Lambda^{q+1} (u u_x) dx + \frac{1}{2}\int_R (\Lambda^q u_x) \Lambda^q \left(u_x^2\right) dx \\ &+ \int_R (\Lambda^q u) \Lambda^q (u u_x) dx - \beta \int_R \Lambda^q u_x \Lambda^q \left[(u_x)^N\right] dx. \end{aligned} \tag{4.9}$$

to Contents, page 1

We will estimate the terms on the right-hand side of (4.9) separately. For the first term, by using the Cauchy-Schwartz inequality and Lemmas 4.1 and 4.2, we have

$$\begin{aligned}\int_R (\Lambda^q u)\Lambda^q(u^m u_x)dx &= \int_R (\Lambda^q u)[\Lambda^q(u^m u_x) - u^m\Lambda^q u_x]dx + \int_R (\Lambda^q u)u^m\Lambda^q u_x dx \\ &\le c\|u\|_{H^q}\left(m\|u\|_{L^\infty}^{m-1}\|u_x\|_{L^\infty}\|u\|_{H^q} + \|u_x\|_{L^\infty}\|u\|_{L^\infty}^{m-1}\|u\|_{H^q}\right) \\ &\quad + \frac{1}{2}\|u\|_{L^\infty}^{m-1}\|u_x\|_{L^\infty}\|\Lambda^q u\|_{L^2}^2 \\ &\le c\|u\|_{H^q}^2\|u\|_{L^\infty}^{m-1}\|u_x\|_{L^\infty}.\end{aligned} \tag{4.10}$$

Using the above estimate to the second term yields

$$\int_R \left(\Lambda^{q+1}u\right)\Lambda^{q+1}(uu_x)dx \le c\|u\|_{H^{q+1}}^2\|u_x\|_{L^\infty}. \tag{4.11}$$

For the third term, using the Cauchy-Schwartz inequality and Lemma 4.1, we obtain

$$\begin{aligned}\int_R (\Lambda^q u_x)\Lambda^q\left(u_x^2\right)dx &\le \|\Lambda^q u_x\|_{L^2}\left\|\Lambda^q\left(u_x^2\right)\right\|_{L^2} \\ &\le c\|u\|_{H^{q+1}}(\|u_x\|_{L^\infty}\|u_x\|_{H^q} + \|u_x\|_{L^\infty}\|u_x\|_{H^q}) \\ &\le c\|u\|_{H^{q+1}}^2\|u_x\|_{L^\infty}.\end{aligned} \tag{4.12}$$

For the last term in (4.9), using Lemma 4.1 repeatedly results in

$$\begin{aligned}\left|\int_R (\Lambda^q u_x)\Lambda^q(u_x)^N dx\right| &\le \|u_x\|_{H^q}\left\|u_x^N\right\|_{H^q} \\ &\le c\|u\|_{H^{q+1}}^2\|u_x\|_{L^\infty}^{N-1}.\end{aligned} \tag{4.13}$$

It follows from (4.9) to (4.13) that there exists a constant c depending only on m, N and the coefficients of (1.2) such that

$$\frac{1}{2}\frac{d}{dt}\int_R \left[(\Lambda^q u)^2 + (\Lambda^q u_x)^2\right]dx \le c\|u_x\|_{L^\infty}\|u\|_{H^{q+1}}^2\left(1 + \|u\|_{L^\infty}^{m-1}\right) + c\|u\|_{H^{q+1}}^2\|u_x\|_{L^\infty}^{N-1}. \tag{4.14}$$

Integrating both sides of the above inequality with respect to t results in inequality (4.6).

To estimate the norm of u_t, we apply the operator $(1-\partial_x^2)^{-1}$ to both sides of the first equation of system (2.3) to obtain the equation

$$u_t = \left(1-\partial_x^2\right)^{-1}\left[-2ku_x + \partial_x\left(-\frac{a}{m+1}u^{m+1} + \frac{1}{2}\partial_x^2\left(u^2\right) - \frac{1}{2}u_x^2\right) + \beta\partial_x\left[(u_x)^N\right]\right]. \tag{4.15}$$

to Contents, page 1

Applying $(\Lambda^q u_t)\Lambda^q$ to both sides of (4.15) for $q \in (0, s-1]$ gives rise to

$$\int_R (\Lambda^q u_t)^2 dx = \int_R (\Lambda^q u_t)\Lambda^{q-2}\left[\partial_x\left(-2ku - \frac{a}{m+1}u^{m+1} + \frac{1}{2}\partial_x^2\left(u^2\right) - \frac{1}{2}u_x^2\right) + \beta\partial_x\left[(u_x)^N\right]\right]d\tau. \tag{4.16}$$

For the right-hand side of (4.16), we have

$$\begin{aligned}
&\int_R (\Lambda^q u_t)\Lambda^{q-2}(-2ku_x)dx \le c\|u_t\|_{H^q}\|u\|_{H^q},\\
&\int_R (\Lambda^q u_t)\left(1-\partial_x^2\right)^{-1}\Lambda^q\partial_x\left(-\frac{a}{m+1}u^{m+1} - \frac{1}{2}u_x^2\right)dx\\
&\quad \le c\|u_t\|_{H^q}\left(\int_R \left(1+\xi^2\right)^{q-1} \times \left[\int_R\left[-\frac{a}{m+1}\widehat{u^m}(\xi-\eta)\hat{u}(\eta) - \frac{1}{2}\widehat{u_x}(\xi-\eta)\widehat{u_x}(\eta)\right]d\eta\right]^2\right)^{1/2}\\
&\quad \le c\|u_t\|_{H^q}\|u\|_{H^1}\|u\|_{H^{q+1}}\left(1+\|u\|_{L^\infty}^{m-1}\right).
\end{aligned} \tag{4.17}$$

Since

$$\int (\Lambda^q u_t)\left(1-\partial_x^2\right)^{-1}\Lambda^q\partial_x^2(uu_x)dx = -\int (\Lambda^q u_t)\Lambda^q(uu_x)dx + \int (\Lambda^q u_t)\left(1-\partial_x^2\right)^{-1}\Lambda^q(uu_x)dx, \tag{4.18}$$

using Lemma 4.1, $\|uu_x\|_{H^q} \le c\|(u^2)_x\|_{H^q} \le c\|u\|_{L^\infty}\|u\|_{H^{q+1}}$ and $\|u\|_{L^\infty} \le \|u\|_{H^1}$, we have

$$\begin{aligned}
\int (\Lambda^q u_t)\Lambda^q(uu_x)dx &\le c\|u_t\|_{H^q}\|uu_x\|_{H^q}\\
&\le c\|u_t\|_{H^q}\|u\|_{H^1}\|u\|_{H^{q+1}},\\
\int (\Lambda^q u_t)\left(1-\partial_x^2\right)^{-1}\Lambda^q(uu_x)dx &\le c\|u_t\|_{H^q}\|u\|_{H^1}\|u\|_{H^{q+1}}.
\end{aligned} \tag{4.19}$$

Using the Cauchy-Schwartz inequality and Lemma 4.1 yields

$$\left|\int_R (\Lambda^q u_t)\left(1-\partial_x^2\right)^{-1}\Lambda^q\partial_x\left(u_x^N\right)dx\right| \le c\|u_t\|_{H^q}\|u_x\|_{L^\infty}^{N-1}\|u\|_{H^{q+1}}. \tag{4.20}$$

Applying (4.17)–(4.20) into (4.16) yields the inequality

$$\|u_t\|_{H^q} \le c\|u\|_{H^{q+1}}\left(1+\left(1+\|u\|_{L^\infty}^{m-1}\right)\|u\|_{H^1} + \|u_x\|_{L^\infty}^{N-1}\right). \tag{4.21}$$

This completes the proof of Lemma 4.3. □

to Contents, page 1

Defining

$$\phi(x) = \begin{cases} e^{1/(x^2-1)}, & |x| < 1, \\ 0, & |x| \geq 1, \end{cases} \tag{4.22}$$

and setting $\phi_\varepsilon(x) = \varepsilon^{-1/4}\phi(\varepsilon^{-1/4}x)$ with $0 < \varepsilon < 1/4$ and $u_{\varepsilon 0} = \phi_\varepsilon \star u_0$, we know that $u_{\varepsilon 0} \in C^\infty$ for any $u_0 \in H^s(R)$ and $s > 0$.

It follows from Theorem 2.1 that for each ε the Cauchy problem

$$\begin{gathered} u_t - u_{txx} = \partial_x\left(-2ku - \frac{a}{m+1}u^{m+1}\right) + \frac{1}{2}\partial_x^3\left(u^2\right) - \frac{1}{2}\partial_x\left(u_x^2\right) + \beta\partial_x\left[(u_x)^N\right], \\ u(0,x) = u_{\varepsilon 0}(x), \quad x \in R, \end{gathered} \tag{4.23}$$

has a unique solution $u_\varepsilon(t,x) \in C^\infty([0,T); H^\infty)$.

Lemma 4.4. *Under the assumptions of problem* (4.23)*, the following estimates hold for any ε with $0 < \varepsilon < 1/4$ and $s > 0$*

$$\begin{gathered} \|u_{\varepsilon 0x}\|_{L^\infty} \leq c_1\|u_{0x}\|_{L^\infty}, \\ \|u_{\varepsilon 0}\|_{H^q} \leq c_1, \quad \text{if } q \leq s, \\ \|u_{\varepsilon 0}\|_{H^q} \leq c_1\varepsilon^{(s-q)/4}, \quad \text{if } q > s, \\ \|u_{\varepsilon 0} - u_0\|_{H^q} \leq c_1\varepsilon^{(s-q)/4}, \quad \text{if } q \leq s, \\ \|u_{\varepsilon 0} - u_0\|_{H^s} = o(1), \end{gathered} \tag{4.24}$$

where c_1 is a constant independent of ε.

The proof of this Lemma can be found in Lai and Wu [33].

Lemma 4.5. *If $u_0(x) \in H^s(R)$ with $s \in [1, 3/2]$ such that $\|u_{0x}\|_{L^\infty} < \infty$. Let $u_{\varepsilon 0}$ be defined as in system* (4.23)*. Then there exist two positive constants T and c, which are independent of ε, such that the solution u_ε of problem* (4.23) *satisfies $\|u_{\varepsilon x}\|_{L^\infty} \leq c$ for any $t \in [0,T)$.*

Proof. Using notation $u = u_\varepsilon$ and differentiating both sides of the first equation of problem (4.23) or (4.15) with respect to x give rise to

$$\begin{aligned} u_{tx} + \frac{1}{2}\partial_x^2 u^2 - \frac{1}{2}u_x^2 = 2ku &+ \frac{a}{m+1}u^{m+1} - \frac{1}{2}u^2 - \beta u_x^N \\ &- \Lambda^{-2}\left[2ku + \frac{a}{m+1}u^{m+1} - \frac{1}{2}u^2 + \frac{1}{2}u_x^2 - \beta u_x^N\right]. \end{aligned} \tag{4.25}$$

to Contents, page 1

Letting $p > 0$ be an integer and multiplying the above equation by $(u_x)^{2p+1}$ and then integrating the resulting equation with respect to x yield the equality

$$\begin{aligned}\frac{1}{2p+2}\frac{d}{dt}\int_R (u_x)^{2p+2}dx + \frac{p}{2p+2}\int_R (u_x)^{2p+3}dx &= \int_R (u_x)^{2p+1}\left(2ku + \frac{a}{m+1}u^{m+1} - \frac{1}{2}u^2 - \beta u_x^N\right)dx \\ &\quad - \int_R (u_x)^{2p+1}\Lambda^{-2}\left[2ku + \frac{a}{m+1}u^{m+1} - \frac{u^2}{2} + \frac{1}{2}u_x^2 - \beta u_x^N\right]dx.\end{aligned} \tag{4.26}$$

Applying the Hölder's inequality yields

$$\begin{aligned}\frac{1}{2p+2}\frac{d}{dt}\int_R (u_x)^{2p+2}dx \le{}& \Bigg\{|2k|\left(\int_R |u|^{2p+2}dx\right)^{1/(2p+2)} + \frac{a}{m+1}\left(\int_R \left|u^{m+1}\right|^{2p+2}dx\right)^{1/(2p+2)} \\ &+ \frac{1}{2}\left(\int_R \left|u^2\right|^{2p+2}dx\right)^{1/(2p+2)} + \beta\left(\int_R \left|u_x^N\right|^{2p+2}dx\right)^{1/(2p+2)} \\ &+ \left(\int_R |G|^{2p+2}dx\right)^{1/(2p+2)}\Bigg\}\left(\int_R |u_x|^{2p+2}dx\right)^{(2p+1)/(2p+2)} \\ &+ \frac{p}{2p+2}\|u_x\|_{L^\infty}\int_R |u_x|^{2p+2}dx,\end{aligned} \tag{4.27}$$

or

$$\begin{aligned}\frac{d}{dt}\left(\int_R (u_x)^{2p+2}dx\right)^{1/(2p+2)} \le{}& |2k|\left(\int_R |u|^{2p+2}dx\right)^{1/(2p+2)} + \frac{a}{m+1}\left(\int_R \left|u^{m+1}\right|^{2p+2}dx\right)^{1/(2p+2)} \\ &+ \frac{1}{2}\left(\int_R \left|u^2\right|^{2p+2}dx\right)^{1/(2p+2)} + \beta\left(\int_R \left|u_x^N\right|^{2p+2}dx\right)^{1/(2p+2)} \\ &+ \left(\int_R |G|^{2p+2}dx\right)^{1/(2p+2)} + \frac{p}{2p+2}\|u_x\|_{L^\infty}\left(\int_R |u_x|^{2p+2}dx\right)^{1/(2p+2)},\end{aligned} \tag{4.28}$$

where

$$G = \Lambda^{-2}\left[2ku + \frac{a}{m+1}u^{m+1} - \frac{u^2}{2} + \frac{1}{2}u_x^2 - \beta u_x^N\right]. \tag{4.29}$$

to Contents, page 1

Since $\|f\|_{L^p} \to \|f\|_{L^\infty}$ as $p \to \infty$ for any $f \in L^\infty \bigcap L^2$, integrating both sides of the inequality (4.28) with respect to t and taking the limit as $p \to \infty$ result in the estimate

$$\|u_x\|_{L^\infty} \le \|u_{0x}\|_{L^\infty} + \int_0^t c\left[\left(\|u\|_{L^\infty} + \left\|u^2\right\|_{L^\infty} + \left\|u^{m+1}\right\|_{L^\infty} + \beta\|u_x\|_{L^\infty}^N + \|G\|_{L^\infty}\right) + \frac{1}{2}\|u_x\|_{L^\infty}^2\right] d\tau. \tag{4.30}$$

Using the algebra property of $H^{s_0}(R)$ with $s_0 > 1/2$ yields ($\|u_\varepsilon\|_{H^{(1/2)+}}$ means that there exists a sufficiently small $\delta > 0$ such that $\|u_\varepsilon\|_{(1/2)+} = \|u_\varepsilon\|_{H^{1/2+\delta}}$)

$$\begin{aligned}
\|G\|_{L^\infty} &\le c\|G\|_{H^{(1/2)+}} \\
&\le c\left\|\Lambda^{-2}\left[2ku + \frac{a}{m+1}u^{m+1} - \frac{u^2}{2} + \frac{1}{2}u_x^2 - \beta u_x^N\right]\right\|_{H^{(1/2)+}} \\
&\le c\left(\|u\|_{H^1} + \|u\|_{H^1}^2 + \|u\|_{H^1}^{m+1} + \left\|\Lambda^{-2}(u_x^2)\right\|_{H^{(1/2)+}} + \left\|\Lambda^{-2}(u_x^N)\right\|_{H^{(1/2)+}}\right) \\
&\le c\left(\|u\|_{H^1} + \|u\|_{H^1}^2 + \|u\|_{H^1}^{m+1} + \left\|u_x^2\right\|_{H^0} + \left\|u_x^N\right\|_{H^0}\right) \\
&\le c\left(\|u\|_{H^1} + \|u\|_{H^1}^2 + \|u\|_{H^1}^{m+1} + \|u_x\|_{L^\infty}\|u\|_{H^1} + \|u_x\|_{L^\infty}^{N-1}\|u\|_{H^1}\right) \\
&\le ce^{c\int_0^t \|u_x\|_{L^\infty}^{N-1} d\tau}\left(1 + \|u_x\|_{L^\infty} + \|u_x\|_{L^\infty}^{N-1}\right),
\end{aligned} \tag{4.31}$$

in which Lemma 4.3 is used. Therefore, we get

$$\int_0^t \|G\|_{L^\infty} d\tau \le c\int_0^t e^{c\int_0^\tau \|u_x\|_{L^\infty}^{N-1} d\xi}\left(1 + \|u_x\|_{L^\infty} + \|u_x\|_{L^\infty}^{N-1}\right) d\tau. \tag{4.32}$$

From (4.30) and (4.32), one has

$$\begin{aligned}
\|u_x\|_{L^\infty} \le \|u_{0x}\|_{L^\infty} + c\int_0^t \Big[&\|u_x\|_{L^\infty}^2 + \|u_x\|_{L^\infty}^N + e^{c\int_0^t \|u_x\|_{L^\infty}^{N-1} d\tau} \\
&+ e^{c\int_0^\tau \|u_x\|_{L^\infty}^{N-1} d\xi}\left(1 + \|u_x\|_{L^\infty} + \|u_x\|_{L^\infty}^{N-1}\right)\Big] d\tau.
\end{aligned} \tag{4.33}$$

From Lemma 4.4, it follows from the contraction mapping principle that there is a $T > 0$ such that the equation

$$\begin{aligned}
\|W\|_{L^\infty} = \|u_{0x}\|_{L^\infty} + c\int_0^t \Big[&\|W\|_{L^\infty}^2 + \|W\|_{L^\infty}^N + e^{c\int_0^t \|W\|_{L^\infty}^{N-1} d\tau} \\
&+ e^{c\int_0^\tau \|W\|_{L^\infty}^{N-1} d\xi}\left(1 + \|W\|_{L^\infty} + \|W\|_{L^\infty}^{N-1}\right)\Big] d\tau
\end{aligned} \tag{4.34}$$

has a unique solution $W \in C[0,T]$. Using the Theorem presented at page 51 in [25] or Theorem 2 in Section 1.1 presented in [37] yields that there are constants $T > 0$ and $c > 0$

to Contents, page 1

independent of ε such that $\|u_x\|_{L^\infty} \leq W(t)$ for arbitrary $t \in [0,T]$, which leads to the conclusion of Lemma 4.5.

Using Lemmas 4.3 and 4.5, notation $u_\varepsilon = u$ and Gronwall's inequality results in the inequalities

$$\begin{aligned} \|u_\varepsilon\|_{H^q} &\leq C_T e^{C_T}, \\ \|u_{\varepsilon t}\|_{H^r} &\leq C_T e^{C_T}, \end{aligned} \tag{4.35}$$

where $q \in (0,s]$, $r \in (0,s-1]$ and C_T depends on T. It follows from Aubin's compactness theorem that there is a subsequence of $\{u_\varepsilon\}$, denoted by $\{u_{\varepsilon_n}\}$, such that $\{u_{\varepsilon_n}\}$ and their temporal derivatives $\{u_{\varepsilon_n t}\}$ are weakly convergent to a function $u(t,x)$ and its derivative u_t in $L^2([0,T],H^s)$ and $L^2([0,T],H^{s-1})$, respectively. Moreover, for any real number $R_1 > 0$, $\{u_{\varepsilon_n}\}$ is convergent to the function u strongly in the space $L^2([0,T],H^q(-R_1,R_1))$ for $q \in [0,s)$ and $\{u_{\varepsilon_n t}\}$ converges to u_t strongly in the space $L^2([0,T],\ H^r(-R_1,R_1))$ for $r \in [0,s-1]$. Thus, we can prove the existence of a weak solution to (2.2). □

Proof of Theorem 2.2. From Lemma 4.5, we know that $\{u_{\varepsilon_n x}\}$ $(\varepsilon_n \to 0)$ is bounded in the space L^∞. Thus, the sequences $\{u_{\varepsilon_n}\}$ and $\{u_{\varepsilon_n x}\}$ are weakly convergent to u and u_x in $L^2[0,T], H^r(-R,R)$ for any $r \in [0,s-1)$, respectively. Therefore, u satisfies the equation

$$\begin{aligned} -\int_0^T \int_R u(g_t - g_{xxt})\,dx\,dt = \int_0^T \int_R \Bigg[&\left(2ku + \frac{a}{m+1}u^{m+1} + \frac{1}{2}\left(u_x^2\right)\right) g_x \\ &- \frac{1}{2}u^2 g_{xxx} - \beta (u_x)^N g_x \Bigg] dx\,dt, \end{aligned} \tag{4.36}$$

with $u(0,x) = u_0(x)$ and $g \in C_0^\infty$. Since $X = L^1([0,T]\times R)$ is a separable Banach space and $\{u_{\varepsilon_n x}\}$ is a bounded sequence in the dual space $X^* = L^\infty([0,T]\times R)$ of X, there exists a subsequence of $\{u_{\varepsilon_n x}\}$, still denoted by $\{u_{\varepsilon_n x}\}$, weakly star convergent to a function v in $L^\infty([0,T]\times R)$. It derives from the $\{u_{\varepsilon_n x}\}$ weakly convergent to u_x in $L^2([0,T]\times R)$ that $u_x = v$ almost everywhere. Thus, we obtain $u_x \in L^\infty([0,T]\times R)$. □

Acknowledgments

The author is very grateful to the reviewers for their helpful and valuable comments, which have led to a meaningful improvement of the paper. This work is supported by the Key Project of Chinese Ministry of Education (109140).

References

[1] R. Camassa and D. D. Holm, "An integrable shallow water equation with peaked solitons," *Physical Review Letters*, vol. 71, no. 11, pp. 1661–1664, 1993.
[2] A. Constantin and D. Lannes, "The hydrodynamical relevance of the Camassa-Holm and Degasperis-Procesi equations," *Archive for Rational Mechanics and Analysis*, vol. 192, no. 1, pp. 165–186, 2009.
[3] R. S. Johnson, "Camassa-Holm, Korteweg-de Vries and related models for water waves," *Journal of Fluid Mechanics*, vol. 455, no. 1, pp. 63–82, 2002.

to Contents, page 1

[4] R. S. Johnson, "The Camassa-Holm equation for water waves moving over a shear flow," *Fluid Dynamics Research*, vol. 33, no. 1-2, pp. 97–111, 2003.
[5] H.-H. Dai, "Model equations for nonlinear dispersive waves in a compressible Mooney-Rivlin rod," *Acta Mechanica*, vol. 127, no. 1–4, pp. 193–207, 1998.
[6] A. Constantin and W. A. Strauss, "Stability of a class of solitary waves in compressible elastic rods," *Physics Letters A*, vol. 270, no. 3-4, pp. 140–148, 2000.
[7] M. Lakshmanan, "Integrable nonlinear wave equations and possible connections to tsunami dynamics," in *Tsunami and Nonlinear Waves*, A. Kundu, Ed., pp. 31–49, Springer, Berlin, Germany, 2007.
[8] A. Constantin and R. S. Johnson, "Propagation of very long water waves, with vorticity, over variable depth, with applications to tsunamis," *Fluid Dynamics Research*, vol. 40, no. 3, pp. 175–211, 2008.
[9] A. Constantin, "The trajectories of particles in Stokes waves," *Inventiones Mathematicae*, vol. 166, no. 3, pp. 523–535, 2006.
[10] A. Constantin and J. Escher, "Particle trajectories in solitary water waves," *Bulletin of American Mathematical Society*, vol. 44, no. 3, pp. 423–431, 2007.
[11] A. Constantin and W. A. Strauss, "Stability of peakons," *Communications on Pure and Applied Mathematics*, vol. 53, no. 5, pp. 603–610, 2000.
[12] A. Constantin and H. P. McKean, "A shallow water equation on the circle," *Communications on Pure and Applied Mathematics*, vol. 52, no. 8, pp. 949–982, 1999.
[13] A. Constantin, "On the inverse spectral problem for the Camassa-Holm equation," *Journal of Functional Analysis*, vol. 155, no. 2, pp. 352–363, 1998.
[14] A. Constantin, V. S. Gerdjikov, and R. I. Ivanov, "Inverse scattering transform for the Camassa-Holm equation," *Inverse Problems*, vol. 22, no. 6, pp. 2197–2207, 2006.
[15] H. P. McKean, "Integrable systems and algebraic curves," in *Global Analysis*, vol. 755 of *Lecture Notes in Mathematics*, pp. 83–200, Springer, Berlin, Germany, 1979.
[16] A. Constantin and B. Kolev, "Geodesic flow on the diffeomorphism group of the circle," *Commentarii Mathematici Helvetici*, vol. 78, no. 4, pp. 787–804, 2003.
[17] G. Misiołek, "A shallow water equation as a geodesic flow on the Bott-Virasoro group," *Journal of Geometry and Physics*, vol. 24, no. 3, pp. 203–208, 1998.
[18] Z. Xin and P. Zhang, "On the weak solutions to a shallow water equation," *Communications on Pure and Applied Mathematics*, vol. 53, no. 11, pp. 1411–1433, 2000.
[19] Z. Xin and P. Zhang, "On the uniqueness and large time behavior of the weak solutions to a shallow water equation," *Communications in Partial Differential Equations*, vol. 27, no. 9-10, pp. 1815–1844, 2002.
[20] G. M. Coclite, H. Holden, and K. H. Karlsen, "Global weak solutions to a generalized hyperelastic-rod wave equation," *SIAM Journal on Mathematical Analysis*, vol. 37, no. 4, pp. 1044–1069, 2005.
[21] A. Bressan and A. Constantin, "Global conservative solutions of the Camassa-Holm equation," *Archive for Rational Mechanics and Analysis*, vol. 183, no. 2, pp. 215–239, 2007.
[22] A. Bressan and A. Constantin, "Global dissipative solutions of the Camassa-Holm equation," *Analysis and Applications*, vol. 5, no. 1, pp. 1–27, 2007.
[23] H. Holden and X. Raynaud, "Dissipative solutions for the Camassa-Holm equation," *Discrete and Continuous Dynamical Systems A*, vol. 24, no. 4, pp. 1047–1112, 2009.
[24] H. Holden and X. Raynaud, "Global conservative solutions of the Camassa-Holm equation—a Lagrangian point of view," *Communications in Partial Differential Equations*, vol. 32, no. 10–12, pp. 1511–1549, 2007.
[25] Y. A. Li and P. J. Olver, "Well-posedness and blow-up solutions for an integrable nonlinearly dispersive model wave equation," *Journal of Differential Equations*, vol. 162, no. 1, pp. 27–63, 2000.
[26] A. Constantin and J. Escher, "Wave breaking for nonlinear nonlocal shallow water equations," *Acta Mathematica*, vol. 181, no. 2, pp. 229–243, 1998.
[27] R. Beals, D. H. Sattinger, and J. Szmigielski, "Multi-peakons and a theorem of Stieltjes," *Inverse Problems*, vol. 15, no. 1, pp. L1–L4, 1999.
[28] D. Henry, "Persistence properties for a family of nonlinear partial differential equations," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 70, no. 4, pp. 1565–1573, 2009.
[29] G. Rodríguez-Blanco, "On the Cauchy problem for the Camassa-Holm equation," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 46, no. 3, pp. 309–327, 2001.
[30] Z. Yin, "On the blow-up scenario for the generalized Camassa-Holm equation," *Communications in Partial Differential Equations*, vol. 29, no. 5-6, pp. 867–877, 2004.
[31] Z. Yin, "On the Cauchy problem for an integrable equation with peakon solutions," *Illinois Journal of Mathematics*, vol. 47, no. 3, pp. 649–666, 2003.

to Contents, page 1

[32] Y. Zhou, "Wave breaking for a periodic shallow water equation," *Journal of Mathematical Analysis and Applications*, vol. 290, no. 2, pp. 591–604, 2004.

[33] S. Lai and Y. Wu, "The local well-posedness and existence of weak solutions for a generalized Camassa-Holm equation," *Journal of Differential Equations*, vol. 248, no. 8, pp. 2038–2063, 2010.

[34] S. Wu and Z. Yin, "Global existence and blow-up phenomena for the weakly dissipative Camassa-Holm equation," *Journal of Differential Equations*, vol. 246, no. 11, pp. 4309–4321, 2009.

[35] T. Kato, "Quasi-linear equations of evolution, with applications to partial differential equations," in *Spectral Theory and Differential Equations*, vol. 448 of *Lecture Notes in Mathematics*, pp. 25–70, Springer, Berlin, Germany, 1975.

[36] T. Kato and G. Ponce, "Commutator estimates and the Euler and Navier-Stokes equations," *Communications on Pure and Applied Mathematics*, vol. 41, no. 7, pp. 891–907, 1988.

[37] W. Walter, *Differential and Integral Inequalities*, vol. 55 of *Ergebnisse der Mathematik und Ihrer Grenzgebiete*, Springer, New York, NY, USA, 1970.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 539026, 36 pages
doi:10.1155/2011/539026

Research Article

Translation Invariant Spaces and Asymptotic Properties of Variational Equations

Adina Luminiţa Sasu and Bogdan Sasu

Department of Mathematics, Faculty of Mathematics and Computer Science, West University of Timişoara, V. Pârvan Boulelvard 4, 300223 Timişoara, Romania

Correspondence should be addressed to Bogdan Sasu, bsasu@math.uvt.ro

Received 11 November 2010; Accepted 12 February 2011

Academic Editor: Josef Diblík

We present a new perspective concerning the study of the asymptotic behavior of variational equations by employing function spaces techniques. We give a complete description of the dichotomous behaviors of the most general case of skew-product flows, without any assumption concerning the flow, the cocycle or the splitting of the state space, our study being based only on the solvability of some associated control systems between certain function spaces. The main results do not only point out new necessary and sufficient conditions for the existence of uniform and exponential dichotomy of skew-product flows, but also provide a clear chart of the connections between the classes of translation invariant function spaces that play the role of the input or output classes with respect to certain control systems. Finally, we emphasize the significance of each underlying hypothesis by illustrative examples and present several interesting applications.

1. Introduction

Starting from a collection of open questions related to the modeling of the equations of mathematical physics in the unified setting of dynamical systems, the study of their qualitative properties became a domain of large interest and with a wide applicability area. In this context, the interaction between the modern methods of pure mathematics and questions arising naturally from mathematical physics created a very active field of research (see [1–18] and the references therein). In recent years, some interesting unsolved problems concerning the long-time behavior of dynamical systems were identified, whose potential results would be of major importance in the process of understanding, clarifying, and solving some of the essential problems belonging to a wide range of scientific domains, among, we mention: fluid mechanics, aeronautics, magnetism, ecology, population dynamics, and so forth. Generally, the asymptotic behavior of the solutions of nonlinear evolution equations

to Contents, page 1

arising in mathematical physics can be described in terms of attractors, which are often studied by constructing the skew-product flows of the dynamical processes.

It was natural then to independently consider and analyze the asymptotic behavior of variational systems modeled by skew-product flows (see [3–5, 14–19]). In this framework, two of the most important asymptotic properties are described by uniform dichotomy and exponential dichotomy. Both properties focus on the decomposition of the state space into a direct sum of two closed invariant subspaces such that the solution on these subspaces (uniformly or exponentially) decays backward and forward in time, and the splitting holds at every point of the flow's domain. Precisely, these phenomena naturally lead to the study of the existence of stable and unstable invariant manifolds. It is worth mentioning that starting with the remarkable works of Coppel [20], Daleckii and Krein [21], and Massera and Schäffer [22] the study of the dichotomy had a notable impact on the development of the qualitative theory of dynamical systems (see [1–9, 13, 14, 17, 18, 23]).

A very important step in the infinite-dimensional asymptotic theory of dynamical systems was made by Van Minh et al. in [7] where the authors proposed a unified treatment of the stability, instability, and dichotomy of evolution families on the half-line via input-output techniques. Their paper carried out a beautiful connection between the classical techniques originating in the pioneering works of Perron [11] and Maĭzel [24] and the natural requests imposed by the development of the infinite-dimensional systems theory. They extended the applicability area of the so-called admissibility techniques developed by Massera and Schäffer in [22], from differential equations in infinite-dimensional spaces to general evolutionary processes described by propagators. The authors pointed out that instead of characterizing the behavior of a homogeneous equation in terms of the solvability of the associated inhomogeneous equation (see [20–22]) one may detect the asymptotic properties by analyzing the existence of the solutions of the associated integral system given by the variation of constants formula. These new methods technically moved the central investigation of the qualitative properties into a different sphere, where the study strongly relied on control-type arguments. It is important to mention that the control-type techniques have been also successfully used by Palmer (see [9]) and by Rodrigues and Ruas-Filho (see [13]) in order to formulate characterizations for exponential dichotomy in terms of the Fredholm Alternative. Starting with these papers, the interaction between control theory and the asymptotic theory of dynamical systems became more profound, and the obtained results covered a large variety of open problems (see, e.g., [1, 2, 12, 14–17, 23] and the references therein).

Despite the density of papers devoted to the study of the dichotomy in the past few years and in contrast with the apparent impression that the phenomenon is well understood, a large number of unsolved problems still raise in this topic, most of them concerning the variational case. In the present paper, we will provide a complete answer to such an open question. We start from a natural problem of finding suitable conditions for the existence of uniform dichotomy as well as of exponential dichotomy using control-type methods, emphasizing on the identification of the essential structures involved in such a construction, as the input-output system, the eligible spaces, the interplay between their main properties, the specific lines that make the differences between a necessary and a sufficient condition, and the proper motivation of each underlying condition.

In this paper, we propose an inedit link between the theory of function spaces and the dichotomous behavior of the solutions of infinite dimensional variational systems, which offers a deeper understanding of the subtle mechanisms that govern the control-type approaches in the study of the existence of the invariant stable and unstable manifolds.

to Contents, page 1

We consider the general setting of variational equations described by skew-product flows, and we associate a control system on the real line. Beside obtaining new conditions for the existence of uniform or exponential dichotomy of skew-product flows, the main aim is to clarify the chart of the connections between the classes of translation invariant function spaces that play the role of the input class or of the output class with respect to the associated control system, proposing a merger between the functional methods proceeding from interpolation theory and the qualitative techniques from the asymptotic theory of dynamical systems in infinite dimensional spaces.

We consider the most general case of skew-product flows, without any assumption concerning the flow or the cocycle, without any invertibility property, and we work without assuming any initial splitting of the state space and without imposing any invariance property. Our central aim is to establish the existence of the dichotomous behaviors with all their properties (see Definitions 3.5 and 4.1) based only on the minimal solvability of an associated control system described at every point of the base space by an integral equation on the real line. First, we deduce conditions for the existence of uniform dichotomy of skew-product flows and we discuss the technical consequences implied by the solvability of the associated control system between two appropriate translation invariant spaces. We point out, for the first time, that an adequate solvability on the real line of the associated integral control system (see Definition 3.6) implies both the existence of the uniform dichotomy projections as well as their uniform boundedness. Next, the attention focuses on the exponential behavior on the stable and unstable manifold, preserving the solvability concept from the previous section and modifying the properties of the input and the output spaces. Thus, we deduce a clear overview on the representative classes of function spaces which should be considered in the detection of the exponential dichotomy of skew-product flows in terms of the solvability of associated control systems on the real line. The obtained results provide not only new necessary and sufficient conditions for exponential dichotomy, but also a complete diagram of the specific delimitations between the classes of function spaces which may be considered in the study of the exponential dichotomy compared with those from the uniform dichotomy case. Moreover, we point out which are the specific properties of the underlying spaces which make a difference between the sufficient hypotheses and the necessary conditions for the existence of exponential dichotomy of skew-product flows. Finally, we motivate our techniques by illustrative examples and present several interesting applications of the main theorems which generalize the input-output type results of previous research in this topic, among, we mention the well-known theorems due to Perron [11], Daleckii and Krein [21], Massera and Schäffer [22], Van Minh et al. [7], and so forth.

2. Banach Function Spaces: Basic Notations and Preliminaries

In this section, for the sake of clarity, we recall several definitions and properties of Banach function spaces, and, also, we establish the notations that will be used throughout the paper.

Let $\mathbb{R}$ denote the set of real numbers, let $\mathbb{R}_+ = \{t \in \mathbb{R} : t \geq 0\}$, and let $\mathbb{R}_- = \{t \in \mathbb{R} : t \leq 0\}$. For every $A \subset \mathbb{R}$, χ_A denotes the characteristic function of the set A. Let $\mathcal{M}(\mathbb{R}, \mathbb{R})$ be the linear space of all Lebesgue measurable functions $u : \mathbb{R} \to \mathbb{R}$ identifying the functions which are equal almost everywhere.

Definition 2.1. A linear subspace $B \subset \mathcal{M}(\mathbb{R}, \mathbb{R})$ is called *normed function space* if there is a mapping $|\cdot|_B : B \to \mathbb{R}_+$ such that the following properties hold:

to Contents, page 1

(i) $|u|_B = 0$ if and only if $u = 0$ a.e.;

(ii) $|\alpha u|_B = |\alpha||u|_B$, for all $(\alpha, u) \in \mathbb{R} \times B$;

(iii) $|u + v|_B \leq |u|_B + |v|_B$, for all $u, v \in B$;

(iv) if $|u(t)| \leq |v(t)|$ a.e. $t \in \mathbb{R}$ and $v \in B$, then $u \in B$ and $|u|_B \leq |v|_B$.

If $(B, |\cdot|_B)$ is complete, then B is called a *Banach function space*.

Remark 2.2. If $(B, |\cdot|_B)$ is a Banach function space and $u \in B$, then also $|u(\cdot)| \in B$.

Definition 2.3. A Banach function space $(B, |\cdot|_B)$ is said to be *invariant under translations* if for every $(u, t) \in B \times \mathbb{R}$ the function $u_t : \mathbb{R} \to \mathbb{R}$, $u_t(s) = u(s - t)$ belongs to B and $|u_t|_B = |u|_B$.

Let $\mathcal{C}_c(\mathbb{R}, \mathbb{R})$ be the linear space of all continuous functions $v : \mathbb{R} \to \mathbb{R}$ with compact support. We denote by $\mathcal{T}(\mathbb{R})$ the class of all Banach function spaces B which are invariant under translations, $\mathcal{C}_c(\mathbb{R}, \mathbb{R}) \subset B$ and

(i) for every $t > 0$ there is $c(t) > 0$ such that $\int_0^t |u(\tau)| d\tau \leq c(t)|u|_B$, for all $u \in B$;

(ii) if $B \setminus L^1(\mathbb{R}, \mathbb{R}) \neq \emptyset$, then there is a continuous function $\gamma \in B \setminus L^1(\mathbb{R}, \mathbb{R})$.

Remark 2.4. Let $B \in \mathcal{T}(\mathbb{R})$. Then, the following properties hold:

(i) if $J \subset \mathbb{R}$ is a bounded interval, then $\chi_J \in B$.

(ii) if $u_n \to u$ in B, then there is a subsequence $(u_{k_n}) \subset (u_n)$ which converges to u a.e. (see, e.g., [25]).

Remark 2.5. Let $B \in \mathcal{T}(\mathbb{R})$. If $\nu > 0$ and $e_\nu : \mathbb{R} \to \mathbb{R}$ is defined by

$$e_\nu(t) = \begin{cases} e^{-\nu t}, & t \geq 0, \\ 0, & t < 0, \end{cases} \tag{2.1}$$

then it is easy to see that

$$e_\nu(t) = \sum_{n=0}^{\infty} e^{-\nu t} \chi_{[n,n+1)}(t) \leq \sum_{n=0}^{\infty} e^{-\nu n} \chi_{[n,n+1)}(t), \quad \forall t \in \mathbb{R}. \tag{2.2}$$

It follows that $e_\nu \in B$ and $|e_\nu|_B \leq |\chi_{[0,1)}|_B / (1 - e^{-\nu})$.

Example 2.6. (i) If $p \in [1, \infty)$, then $L^p(\mathbb{R}, \mathbb{R}) = \{u \in \mathcal{M}(\mathbb{R}, \mathbb{R}) : \int_{\mathbb{R}} |u(t)|^p dt < \infty\}$, with respect to the norm $\|u\|_p = (\int_{\mathbb{R}} |u(t)|^p dt)^{1/p}$, is a Banach function space which belongs to $\mathcal{T}(\mathbb{R})$.

(ii) The linear space $L^\infty(\mathbb{R}, \mathbb{R})$ of all measurable essentially bounded functions $u : \mathbb{R} \to \mathbb{R}$ with respect to the norm $\|u\|_\infty = \text{ess sup}_{t \in \mathbb{R}} |u(t)|$ is a Banach function space which belongs to $\mathcal{T}(\mathbb{R})$.

Example 2.7 (Orlicz spaces). Let $\varphi : \mathbb{R}_+ \to \overline{\mathbb{R}}_+$ be a nondecreasing left continuous function which is not identically 0 or ∞ on $(0, \infty)$, and let $Y_\varphi(t) := \int_0^t \varphi(s) ds$. If $u \in \mathcal{M}(\mathbb{R}, \mathbb{R})$ let

$$M_\varphi(u) := \int_{\mathbb{R}} Y_\varphi(|u(s)|) ds. \tag{2.3}$$

to Contents, page 1

The linear space $O_\varphi(\mathbb{R},\mathbb{R}) := \{u \in \mathcal{M}(\mathbb{R},\mathbb{R}) : \exists k > 0$ such that $M_\varphi(ku) < \infty\}$, with respect to the norm

$$|u|_\varphi := \inf\left\{k > 0 : M_\varphi\left(\frac{u}{k}\right) \leq 1\right\}, \tag{2.4}$$

is a Banach function space called the *Orlicz space* associated to φ. It is easy to see that $O_\varphi(\mathbb{R},\mathbb{R})$ is invariant under translations.

Remark 2.8. A remarkable example of Orlicz space is represented by $L^p(\mathbb{R},\mathbb{R})$, for every $p \in [1,\infty]$. This can be obtained for $\varphi(t) = pt^{p-1}$, if $p \in [1,\infty)$ and for

$$\varphi(t) = \begin{cases} 0, & t \in [0,1], \\ \infty, & t > 1, \end{cases} \quad \text{if } p = \infty. \tag{2.5}$$

Lemma 2.9. *If $\varphi(1) < \infty$, then $O_\varphi(\mathbb{R},\mathbb{R}) \in \mathcal{T}(\mathbb{R})$.*

Proof. Let $v \in C_c(\mathbb{R},\mathbb{R})$. Then, there are $a,\ b \in \mathbb{R}, a < b$ such that $v(t) = 0$, for all $t \in \mathbb{R} \setminus (a,b)$. Since v is continuous on $[a,b]$, there is $M > 0$ such that $|v(t)| \leq M$, for all $t \in [a,b]$. Then, we have that

$$|v(t)| \leq M\chi_{[a,b]}(t), \quad \forall t \in \mathbb{R}. \tag{2.6}$$

We observe that

$$M_\varphi(\chi_{[a,b]}) = \int_{\mathbb{R}} Y_\varphi(\chi_{[a,b]}(\tau))d\tau = (b-a)Y_\varphi(1) \leq (b-a)\varphi(1) < \infty. \tag{2.7}$$

This implies that $\chi_{[a,b]} \in O_\varphi(\mathbb{R},\mathbb{R})$. Using (2.6), we deduce that $v \in O_\varphi(\mathbb{R},\mathbb{R})$. So, $C_c(\mathbb{R},\mathbb{R}) \subset O_\varphi(\mathbb{R},\mathbb{R})$.

Since Y_φ is nondecreasing with $\lim_{t\to\infty} Y_\varphi(t) = \infty$, there is $q > 0$ such that $Y_\varphi(t) > 1$, for all $t \geq q$.

Let $t \geq 1$ and let $u \in O_\varphi(\mathbb{R},\mathbb{R}) \setminus \{0\}$. Taking into account that Y_φ is a convex function and using Jensen's inequality (see, e.g., [26]), we deduce that

$$Y_\varphi\left(\frac{1}{t}\int_0^t \frac{|u(\tau)|}{|u|_\varphi}d\tau\right) \leq \frac{1}{t}\int_0^t Y_\varphi\left(\frac{|u(\tau)|}{|u|_\varphi}\right)d\tau \leq M_\varphi\left(\frac{u}{|u|_\varphi}\right) \leq 1. \tag{2.8}$$

This implies that

$$\frac{1}{t}\int_0^t \frac{|u(\tau)|}{|u|_\varphi}d\tau \leq q, \quad \forall t \geq 1. \tag{2.9}$$

to Contents, page 1

In addition, using (2.9), we have that

$$\int_0^t |u(\tau)|d\tau \leq \int_0^1 |u(\tau)|d\tau \leq q|u|_\varphi, \quad \forall t \in [0,1). \tag{2.10}$$

Taking $c : (0,\infty) \to (0,\infty)$, $c(t) = \max\{qt, q\}$, from relations (2.9) and (2.10), it follows that

$$\int_0^t |u(\tau)|d\tau \leq c(t)|u|_\varphi, \quad \forall t \geq 0. \tag{2.11}$$

Since the function c does not depend on u, we obtain that $O_\varphi(\mathbb{R},\mathbb{R}) \in \mathcal{T}(\mathbb{R})$. □

Example 2.10. If $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ defined by $\varphi(0) = 0$, $\varphi(t) = 1$, for $t \in (0,1]$ and $\varphi(t) = e^{t-1}$, for $t > 1$, then according to Lemma 2.9 we have that the Orlicz space $O_\varphi(\mathbb{R},\mathbb{R}) \in \mathcal{T}(\mathbb{R})$. Moreover, it is easy to see that $O_\varphi(\mathbb{R},\mathbb{R})$ is a proper subspace of $L^1(\mathbb{R},\mathbb{R})$.

Example 2.11. Let $p \in [1,\infty)$ and let $M^p(\mathbb{R},\mathbb{R})$ be the linear space of all $u \in \mathcal{M}(\mathbb{R},\mathbb{R})$ with $\sup_{t\in\mathbb{R}} \int_t^{t+1} |u(s)|^p ds < \infty$. With respect to the norm

$$\|u\|_{M^p} := \sup_{t\in\mathbb{R}} \left(\int_t^{t+1} |u(s)|^p ds \right)^{1/p}, \tag{2.12}$$

this is a Banach function space which belongs to $\mathcal{T}(\mathbb{R})$.

Remark 2.12. If $B \in \mathcal{T}(\mathbb{R})$, then $B \subset M^1(\mathbb{R},\mathbb{R})$.

Indeed, let $c(1) > 0$ be such that $\int_0^1 |u(\tau)|d\tau \leq c(1)|u|_B$, for all $u \in B$. If $u \in B$ we observe that

$$\int_t^{t+1} |u(\tau)|d\tau = \int_0^1 |u_t(\xi)|d\xi \leq c(1)|u_t|_B = c(1)|u|_B, \quad \forall t \in \mathbb{R}, \tag{2.13}$$

so $u \in M^1(\mathbb{R},\mathbb{R})$.

In what follows, we will introduce three remarkable subclasses of $\mathcal{T}(\mathbb{R})$, which will have an essential role in the study of the existence of dichotomy from the next sections. To do this, we first need the following.

Definition 2.13. Let $B \in \mathcal{T}(\mathbb{R})$. The mapping $F_B : (0,\infty) \to \mathbb{R}_+$, $F_B(t) = |\chi_{[0,t)}|_B$ is called *the fundamental function* of the space B.

Remark 2.14. If $B \in \mathcal{T}(\mathbb{R})$, then the fundamental function F_B is nondecreasing.

Notation 1. We denote by $Q(\mathbb{R})$ the class of all Banach function spaces $B \in \mathcal{T}(\mathbb{R})$ with the property that $\sup_{t>0} F_B(t) = \infty$.

to Contents, page 1

Lemma 2.15. *If $\varphi(t) \in (0, \infty)$, for all $t > 0$, then $O_\varphi(\mathbb{R}, \mathbb{R}) \in \mathcal{Q}(\mathbb{R})$.*

Proof. It is easy to see that Y_φ is strictly increasing, continuous with $Y_\varphi(0) = 0$ and $Y_\varphi(t) \geq (t-1)\varphi(1)$, for all $t > 1$, so $\lim_{t\to\infty} Y_\varphi(t) = \infty$. Hence, Y_φ is bijective.

Let $t > 0$. Since

$$M_\varphi\left(\frac{1}{k}\chi_{[0,t)}\right) = tY_\varphi\left(\frac{1}{k}\right), \quad \forall k > 0, \tag{2.14}$$

it follows that $M_\varphi((1/k)\chi_{[0,t)}) \leq 1$ if and only if $1/Y_\varphi^{-1}(1/t) \leq k$. This implies that

$$F_{O_\varphi(\mathbb{R},\mathbb{R})}(t) = \frac{1}{Y_\varphi^{-1}(1/t)}, \quad \forall t > 0. \tag{2.15}$$

Since $Y_\varphi^{-1}(0) = 0$, from (2.15), we obtain that $O_\varphi(\mathbb{R}, \mathbb{R}) \in \mathcal{Q}(\mathbb{R})$. □

Another distinctive subclass of $\mathcal{T}(\mathbb{R})$ is introduced in the following.

Notation 2. Let $\mathcal{L}(\mathbb{R})$ denote the class of all Banach function spaces $B \in \mathcal{T}(\mathbb{R})$ with the property that $B \setminus L^1(\mathbb{R}, \mathbb{R}) \neq \emptyset$.

Remark 2.16. According to Remark 2.2, we have that if $B \in \mathcal{L}(\mathbb{R})$, then there is a continuous function $\gamma : \mathbb{R} \to \mathbb{R}_+$ such that $\gamma \in B \setminus L^1(\mathbb{R}, \mathbb{R})$.

We will also see, in this paper, that the necessary conditions for the existence of exponential dichotomy should be expressed using another remarkable subclass of $\mathcal{T}(\mathbb{R})$—the rearrangement invariant spaces, see the definitions below.

Definition 2.17. Let $u, v \in \mathcal{M}(\mathbb{R}, \mathbb{R})$. We say that u and v are *equimeasurable* if for every $t > 0$ the sets $\{s \in \mathbb{R} : |u(s)| > t\}$ and $\{s \in \mathbb{R} : |v(s)| > t\}$ have the same measure.

Definition 2.18. A Banach function space $(B, |\cdot|_B)$ is *rearrangement invariant* if for every equimeasurable functions u, v with $u \in B$, we have that $v \in B$ and $|u|_B = |v|_B$.

Notation 3. We denote by $\mathcal{R}(\mathbb{R})$ the class of all Banach function spaces $B \in \mathcal{T}(\mathbb{R})$ which are rearrangement invariant.

Remark 2.19. If $B \in \mathcal{R}(\mathbb{R})$, then B is an interpolation space between $L^1(\mathbb{R},\mathbb{R})$ and $L^\infty(\mathbb{R}, \mathbb{R})$ (see [27, Theorem 2.2, page 106]).

Remark 2.20. The Orlicz spaces are rearrangement invariant (see [27]). Using Lemma 2.9, we deduce that if $\varphi(1) < \infty$, then $O_\varphi(\mathbb{R}, \mathbb{R}) \in \mathcal{R}(\mathbb{R})$.

Lemma 2.21. *Let $B \in \mathcal{R}(\mathbb{R})$ and let $\nu > 0$. Then for every $u \in B$, the functions $\varphi_u, \psi_u : \mathbb{R} \to \mathbb{R}$ defined by*

$$\varphi_u(t) = \int_{-\infty}^{t} e^{-\nu(t-\tau)}u(\tau)d\tau, \qquad \psi_u(t) = \int_{t}^{\infty} e^{-\nu(\tau-t)}u(\tau)d\tau \tag{2.16}$$

to Contents, page 1

belong to B. Moreover, there is $\gamma_{B,\nu} > 0$ which depends only on B and ν such that

$$\left|\varphi_u\right|_B \leq \gamma_{B,\nu}|u|_B, \qquad \left|\psi_u\right|_B \leq \gamma_{B,\nu}|u|_B, \quad \forall u \in B. \tag{2.17}$$

Proof. We consider the operators

$$\begin{aligned} Z : L^{\infty}(\mathbb{R}, \mathbb{R}) \longrightarrow L^{\infty}(\mathbb{R}, \mathbb{R}), \quad (Z(u))(t) = \int_{-\infty}^{t} e^{-\nu(t-\tau)} u(\tau) d\tau, \\ W : L^{\infty}(\mathbb{R}, \mathbb{R}) \longrightarrow L^{\infty}(\mathbb{R}, \mathbb{R}), \quad (W(u))(t) = \int_{t}^{\infty} e^{-\nu(\tau-t)} u(\tau) d\tau. \end{aligned} \tag{2.18}$$

We have that Z and W are correctly defined bounded linear operators. Moreover, the restrictions $Z_| : L^1(\mathbb{R}, \mathbb{R}) \to L^1(\mathbb{R}, \mathbb{R})$ and $W_| : L^1(\mathbb{R}, \mathbb{R}) \to L^1(\mathbb{R}, \mathbb{R})$ are correctly defined and bounded linear operators. Since $B \in \mathcal{R}(\mathbb{R})$, then, from Remark 2.19, we have that B is an interpolation space between $L^1(\mathbb{R},\mathbb{R})$ and $L^{\infty}(\mathbb{R}, \mathbb{R})$. This implies that the restrictions $Z_{|B} : B \to B$ and $W_{|B} : B \to B$ are correctly defined and bounded linear operators. Setting $\gamma_{B,\nu} = \max\{\|Z_{|B}\|, \|W_{|B}\|\}$, the proof is complete. □

Notations

If X is a Banach space, for every Banach function space $B \in \mathcal{T}(\mathbb{R})$, we denote by $B(\mathbb{R}, X)$ the space of all Bochner measurable functions $v : \mathbb{R} \to X$ with the property that the mapping $N_v : \mathbb{R} \to \mathbb{R}_+$, $N_v(t) = \|v(t)\|$ belongs to B. With respect to the norm

$$\|v\|_{B(\mathbb{R},X)} := |N_v|_B, \tag{2.19}$$

$B(\mathbb{R}, X)$ is a Banach space. We also denote by $C_{0,c}(\mathbb{R}, X)$ the linear space of all continuous functions $v : \mathbb{R} \to X$ with compact support contained in $(0, \infty)$. It is easy to see that $C_{0,c}(\mathbb{R}, X) \subset B(\mathbb{R}, X)$, for all $B \in \mathcal{T}(\mathbb{R})$.

3. Uniform Dichotomy for Skew-Product Flows

In this section, we start our investigation by studying the existence of by the upper and lower uniform boundedness of the solution in a uniform way on certain complemented subspaces. We will employ a control-type technique and we will show that the use of the function spaces, from the class $\mathcal{T}(\mathbb{R})$ introduced in the previous section, provides several interesting conclusions concerning the qualitative behavior of the solutions of variational equations.

Let X be a real or complex Banach space and let I_d denote the identity operator on X. The norm on X and on $\mathcal{B}(X)$—the Banach algebra of all bounded linear operators on X, will be denoted by $\|\cdot\|$. Let (Θ, d) be a metric space.

Definition 3.1. A continuous mapping $\sigma : \Theta \times \mathbb{R} \to \Theta$ is called a *flow* on Θ if $\sigma(\theta, 0) = \theta$ and $\sigma(\theta, s+t) = \sigma(\sigma(\theta, s), t)$, for all $(\theta, s, t) \in \Theta \times \mathbb{R}^2$.

Definition 3.2. A pair $\pi = (\Phi, \sigma)$ is called a *skew-product flow* on X × Θ if σ is a flow on Θ and the mapping $\Phi : \Theta \times \mathbb{R}_+ \to \mathcal{B}(X)$ called *cocycle*, satisfies the following conditions:

(i) $\Phi(\theta,0)=I_d$ and $\Phi(\theta,t+s)=\Phi(\sigma(\theta,s),t)\Phi(\theta,s)$, for all $(\theta,t,s)\in\Theta\times\mathbb{R}_+^2$;

(ii) there are $M\geq 1$ and $\omega>0$ such that $\|\Phi(\theta,t)\|\leq Me^{\omega t}$, for all $(\theta,t)\in\Theta\times\mathbb{R}_+$;

(iii) for every $(x,\theta)\in X\times\Theta$, the mapping $t\mapsto\Phi(\theta,t)x$ is continuous on $\mathbb{R}_+$.

Example 3.3 (Particular cases). The class described by skew-product flows generalizes the autonomous systems as well as the nonautonomous systems, as the following examples show:

(i) If $\Theta=\mathbb{R}$, then let $\tilde{\sigma}(\theta,t)=\theta+t$ and let $\{U(t,s)\}_{t\geq s}$ be an evolution family on the Banach space X. Setting $\Phi_U(\theta,t):=U(\theta+t,\theta)$, we observe that $\pi_U=(\Phi_U,\tilde{\sigma})$ is a skew-product flow.

(ii) Let $\{T(t)\}_{t\geq 0}$ be a $\mathcal{C}_0$-semigroup on the Banach space X and let Θ be a metric space.

$(\text{ii})_1$ If σ is an arbitrary flow on Θ and $\Phi_T(\theta,t):=T(t)$, then $\pi_T=(\Phi_T,\sigma)$ is a skew-product flow.

$(\text{ii})_2$ Let $\hat{\sigma}:\Theta\times\mathbb{R}\to\Theta$, $\hat{\sigma}(\theta,t)=\theta$ be the projection flow on Θ and let $\{P(\theta)\}_{\theta\in\Theta}\subset\mathcal{B}(X)$ be a uniformly bounded family of projections such that $P(\theta)T(t)=T(t)P(\theta)$, for all $(\theta,t)\in\Theta\times\mathbb{R}_+$. If $\Phi_P(\theta,t):=P(\theta)T(t)$, then $\pi_P=(\Phi_P,\hat{\sigma})$ is a skew-product flow.

Starting with the remarkable work of Foias et al. (see [19]), the qualitative theory of dynamical systems acquired a new perspective concerning the connections between bifurcation theory and the mathematical modeling of nonlinear equations. In [19], the authors proved that classical equations like Navier-Stokes, Taylor-Couette, and Bubnov-Galerkin can be modeled and studied in the unified setting of skew-product flows. In this context, it was pointed out that the skew-product flows often proceed from the linearization of nonlinear equations. Thus, classical examples of skew-product flows arise as operator solutions for variational equations.

Example 3.4 (The variational equation). Let Θ be a locally compact metric space and let σ be a flow on Θ. Let X be a Banach space and let $\{A(\theta):D(A(\theta))\subseteq X\to X:\theta\in\Theta\}$ be a family of densely defined closed operators. We consider the variational equation

$$\dot{x}(t)=A(\sigma(\theta,t))x(t),\quad(\theta,t)\in\Theta\times\mathbb{R}_+.\tag{A}$$

A cocycle $\Phi:\Theta\times\mathbb{R}_+\to\mathcal{B}(X)$ is said to be *a solution of* (A) if for every $\theta\in\Theta$, there is a dense subset $D_\theta\subset D(A(\theta))$ such that for every initial condition $x_\theta\in D_\theta$ the mapping $t\mapsto x(t):=\Phi(\theta,t)x_\theta$ is differentiable on $\mathbb{R}_+$, for every $t\in\mathbb{R}_+ x(t)\in D(A(\sigma(\theta,t)))$ and the mapping $t\mapsto x(t)$ satisfies (A).

An important asymptotic behavior of skew-product flows is described by the uniform dichotomy, which relies on the splitting of the Banach space X at every point $\theta\in\Theta$ into a direct sum of two invariant subspaces such that on the first subspace the trajectory solution is uniformly stable, on the second subspace the restriction of the cocycle is reversible and also the trajectory solution is uniformly unstable on the second subspace. This is given by the following.

to Contents, page 1

Definition 3.5. A skew-product flow $\pi = (\Phi, \sigma)$ is said to be *uniformly dichotomic* if there exist a family of projections $\{P(\theta)\}_{\theta\in\Theta} \subset \mathcal{B}(X)$ and a constant $K \geq 1$ such that the following properties hold:

(i) $\Phi(\theta,t)P(\theta) = P(\sigma(\theta,t))\Phi(\theta,t)$, for all $(\theta,t) \in \Theta \times \mathbb{R}_+$;

(ii) $\|\Phi(\theta,t)x\| \leq K\|x\|$, for all $t \geq 0$, all $x \in \operatorname{Range} P(\theta)$ and all $\theta \in \Theta$;

(iii) the restriction $\Phi(\theta,t)_| : \operatorname{Ker} P(\theta) \to \operatorname{Ker} P(\sigma(\theta,t))$ is an isomorphism, for all $(\theta,t) \in \Theta \times \mathbb{R}_+$;

(iv) $\|\Phi(\theta,t)y\| \geq (1/K)\|y\|$, for all $t \geq 0$, all $y \in \operatorname{Ker} P(\theta)$ and all $\theta \in \Theta$;

(v) $\sup_{\theta\in\Theta}\|P(\theta)\| < \infty$.

In what follows, our main attention will focus on finding suitable conditions for the existence of uniform dichotomy for skew-product flows. To do this, we will introduce an integral control system associated with a skew-product flow such that the input and the output spaces of the system belong to the general class $\mathcal{T}(\mathbb{R})$. We will emphasize that the class $\mathcal{T}(\mathbb{R})$ has an essential role in the study of the dichotomous behavior of variational equations.

Let I, O be two Banach function spaces with $I, O \in \mathcal{T}(\mathbb{R})$. Let $\pi = (\Phi, \sigma)$ be a skew-product flow on $X \times \Theta$. We associate with π the input-output control system $E_\pi = (E_\theta)_{\theta\in\Theta}$, where for every $\theta \in \Theta$

$$f(t) = \Phi(\sigma(\theta,s), t-s)f(s) + \int_s^t \Phi(\sigma(\theta,\tau), t-\tau)v(\tau)d\tau, \quad \forall t \geq s, \tag{E_θ}$$

such that the input function $v \in C_{0,c}(\mathbb{R}, X)$ and the output function $f \in O(\mathbb{R}, X)$.

Definition 3.6. The pair $(O(\mathbb{R},X), I(\mathbb{R},X))$ is said to be *uniformly admissible* for the system (E_π) if there is $L > 0$ such that for every $\theta \in \Theta$, the following properties hold:

(i) for every $v \in C_{0,c}(\mathbb{R}, X)$ there exists $f \in O(\mathbb{R}, X)$ such that the pair (f, v) satisfies (E_θ);

(ii) if $v \in C_{0,c}(\mathbb{R}, X)$ and $f \in O(\mathbb{R}, X)$ are such that the pair (f, v) satisfies (E_θ), then $\|f\|_{O(\mathbb{R},X)} \leq L\|v\|_{I(\mathbb{R},X)}$.

Remark 3.7. (i) According to this admissibility concept, it is sufficient to choose all the input functions from the space $C_{0,c}(\mathbb{R}, X)$, and, thus, we point out that $C_{0,c}(\mathbb{R}, X)$ is in fact *the smaller possible input space* that can be used in the input-output study of the dichotomy.

(ii) It is also interesting to see that the norm estimation from (ii) reflects the presence (and implicitly the structure) of the space $I(\mathbb{R}, X)$. Actually, condition (ii) shows that the norm of each output function in the space $O(\mathbb{R}, X)$ is bounded by the norm of the input function in the space $I(\mathbb{R}, X)$ uniformly with respect to $\theta \in \Theta$.

(iii) In the admissibility concept, there is no need to require the uniqueness of the output function in the property (i), because this follows from condition (ii). Indeed, if the pair $(O(\mathbb{R},X), I(\mathbb{R},X))$ is uniformly admissible for the system (E_π), then from (ii) we deduce that for every $\theta \in \Theta$ and every $v \in C_{0,c}(\mathbb{R}, X)$ there exists *a unique* $f \in O(\mathbb{R}, X)$ such that the pair (f, v) satisfies (E_θ).

In what follows we will analyze the implications of the uniform admissibility of the pair $(O(\mathbb{R},X), I(\mathbb{R},X))$ with $I, O \in \mathcal{T}(\mathbb{R})$ concerning the asymptotic behavior of skew-product

to Contents, page 1

flows. With this purpose we introduce two category of subspaces (stable and unstable) and we will point out their role in the detection of the uniform dichotomy.

For every $(x,\theta) \in X \times \Theta$, we consider the function

$$\lambda_{x,\theta} : \mathbb{R} \longrightarrow X, \quad \lambda_{x,\theta}(t) = \begin{cases} \Phi(\theta,t)x, & t \geq 0, \\ 0, & t < 0, \end{cases} \tag{3.1}$$

called *the trajectory* determined by the vector x and the point $\theta \in \Theta$.

For every $\theta \in \Theta$, we denote by $\mathcal{F}(\theta)$ the linear space of all functions $\varphi : \mathbb{R} \to X$ with the property that

$$\varphi(t) = \Phi(\sigma(\theta,s), t-s)\varphi(s), \quad \forall s \leq t \leq 0. \tag{3.2}$$

For every $\theta \in \Theta$, we consider the *stable subset*

$$\mathcal{S}(\theta) = \{x \in X : \lambda_{x,\theta} \in O(\mathbb{R}, X)\} \tag{3.3}$$

and, respectively, the *unstable subset*

$$\mathcal{U}(\theta) = \{x \in X : \exists\varphi \in O(\mathbb{R}, X) \cap \mathcal{F}(\theta) \text{ with } \varphi(0) = x\}. \tag{3.4}$$

Remark 3.8. It is easy to see that for every $\theta \in \Theta$, $\mathcal{S}(\theta)$, and $\mathcal{U}(\theta)$ are linear subspaces. Therefore, in all what follows, we will refer $\mathcal{S}(\theta)$ as the stable subspace and, respectively, $\mathcal{U}(\theta)$ as the unstable subspace, for each $\theta \in \Theta$.

Proposition 3.9. *For every $(\theta,t) \in \Theta \times \mathbb{R}_+$, the following assertions hold:*

(i) $\Phi(\theta,t)\mathcal{S}(\theta) \subseteq \mathcal{S}(\sigma(\theta,t))$;

(ii) $\Phi(\theta,t)\mathcal{U}(\theta) = \mathcal{U}(\sigma(\theta,t))$.

Proof. The property (i) is immediate. To prove the assertion (ii) let $M, \omega > 0$ be given by Definition 3.2(ii). Let $(\theta,t) \in \Theta \times (0,\infty)$. Let $x \in \mathcal{U}(\theta)$. Then, there is $\varphi \in O(\mathbb{R}, X) \cap \mathcal{F}(\theta)$ with $\varphi(0) = x$. We set $y = \Phi(\theta,t)x$, and we consider

$$\psi : \mathbb{R} \longrightarrow X, \quad \psi(s) = \begin{cases} 0, & s > t, \\ \Phi(\theta,s)x, & s \in [0,t], \\ \varphi(s), & s < 0. \end{cases} \tag{3.5}$$

We observe that $\|\psi(s)\| \leq \|\varphi(s)\| + Me^{\omega t}\chi_{[0,t]}(s)\|x\|$, for all $s \in \mathbb{R}$, and since $\varphi \in O(\mathbb{R}, X)$, we deduce that $\psi \in O(\mathbb{R}, X)$. Using the fact that $\varphi \in \mathcal{F}(\theta)$, we obtain that

$$\psi(s) = \Phi(\sigma(\theta,\tau), s-\tau)\psi(\tau), \quad \forall \tau \leq s \leq t. \tag{3.6}$$

to Contents, page 1

Then, we define the function $\delta : \mathbb{R} \to X$, $\delta(s) = \varphi(s+t)$ and since $O(\mathbb{R}, X)$ is invariant under translations, we deduce that $\delta \in O(\mathbb{R}, X)$. Moreover, from (3.6), it follows that

$$\delta(r) = \Phi(\sigma(\theta, \xi + t), r - \xi)\delta(\xi) = \Phi(\sigma(\sigma(\theta, t), \xi), r - \xi)\delta(\xi), \quad \forall \xi \le r \le 0. \tag{3.7}$$

The relation (3.7) implies that $\delta \in \mathcal{F}(\sigma(\theta, t))$, so $y = \delta(0) \in \mathcal{U}(\sigma(\theta, t))$.

Conversely, let $z \in \mathcal{U}(\sigma(\theta, t))$. Then, there is $h \in \mathcal{F}(\sigma(\theta, t)) \cap O(\mathbb{R}, X)$ with $h(0) = z$. Taking $q : \mathbb{R} \to X$, $q(s) = h(s-t)$, we have that $q \in O(\mathbb{R}, X)$ and

$$q(s) = \Phi(\sigma(\theta, \tau), s - \tau)q(\tau), \quad \forall \tau \le s \le t. \tag{3.8}$$

In particular, for $\tau \le s \le 0$, from (3.8), we deduce that $q \in \mathcal{F}(\theta)$. This implies that $q(0) \in \mathcal{U}(\theta)$. Then, $z = h(0) = q(t) = \Phi(\theta, t)q(0) \in \Phi(\theta, t)\mathcal{U}(\theta)$ and the proof is complete. □

Remark 3.10. From Proposition 3.9(ii), we have that for every $(\theta, t) \in \Theta \times \mathbb{R}_+$ the restriction $\Phi(\theta, t)_| : \mathcal{U}(\theta) \to \mathcal{U}(\sigma(\theta, t))$ is surjective. We also note that according to Proposition 3.9 one may deduce that, the stable subspace and the unstable subspace are candidates for the possible splitting of the main space X required by any dichotomous behavior.

In what follows, we will study the behavior of the cocycle on the stable subspace and also on the unstable subspace and we will deduce several interesting properties of these subspaces in the hypothesis that a pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ of spaces from the class $\mathcal{T}(\mathbb{R})$ is admissible for the control system associated with the skew-product flow.

Theorem 3.11 (The behavior on the stable subspace). *If the pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ is uniformly admissible for the system (E_π), then the following assertions hold:*

(i) *there is $K > 0$ such that $\|\Phi(\theta, t)x\| \le K\|x\|$, for all $t \ge 0$, all $x \in \mathcal{S}(\theta)$ and all $\theta \in \Theta$;*

(ii) *$\mathcal{S}(\theta)$ is a closed linear subspace, for all $\theta \in \Theta$.*

Proof. Let $L > 0$ be given by Definition 3.6 and let $M, \omega > 0$ be given by Definition 3.2. Let $\alpha : \mathbb{R} \to [0, 2]$ be a continuous function with $\operatorname{supp} \alpha \subset (0, 1)$ and $\int_0^1 \alpha(\tau)d\tau = 1$.

(i) Let $\theta \in \Theta$ and let $x \in \mathcal{S}(\theta)$. We consider the functions

$$v : \mathbb{R} \longrightarrow X, \quad v(t) = \alpha(t)\Phi(\theta, t)x,$$
$$f : \mathbb{R} \longrightarrow X, \quad f(t) = \begin{cases} \Phi(\theta, t)x, & t \ge 1, \\ \displaystyle\int_0^t \alpha(\tau)d\tau\, \Phi(\theta, t)x, & t \in [0, 1), \\ 0, & t < 0. \end{cases} \tag{3.9}$$

Then, $v \in C_{0c}(\mathbb{R}, X)$ and

$$\|f(t)\| \le \|\lambda_{x,\theta}(t)\|, \quad \forall t \in \mathbb{R}. \tag{3.10}$$

to Contents, page 1

Since $x \in \mathcal{S}(\theta)$, we have that $\lambda_{x,\theta} \in O(\mathbb{R}, X)$. Then, from (3.10), we obtain that $f \in O(\mathbb{R}, X)$. An easy computation shows that the pair (f, v) satisfies (E_θ). Then,

$$\|f\|_{O(\mathbb{R},X)} \le L\|v\|_{I(\mathbb{R},X)}. \tag{3.11}$$

From $\|v(t)\| \le \alpha(t)Me^{\omega}\|x\|$, for all $t \in \mathbb{R}$, we obtain that $\|v\|_{I(\mathbb{R},X)} \le Me^{\omega}|\alpha|_I\|x\|$.

Let $t \ge 2$. From

$$\|\Phi(\theta,t)x\| \le Me^{\omega}\|\Phi(\theta,s)x\|, \quad \forall s \in [t-1,t), \tag{3.12}$$

it follows that

$$\|\Phi(\theta,t)x\|\chi_{[t-1,t)}(s) \le Me^{\omega}\|f(s)\|, \quad \forall s \in \mathbb{R}. \tag{3.13}$$

Since O is invariant under translations, we deduce that

$$\|\Phi(\theta,t)x\|F_O(1) \le Me^{\omega}\|f\|_{O(\mathbb{R},X)}. \tag{3.14}$$

Using relations (3.11) and (3.14), we have that

$$\|\Phi(\theta,t)x\| \le M^2e^{2\omega}\frac{L|\alpha|_I}{F_O(1)}\|x\|, \quad \forall t \ge 2. \tag{3.15}$$

Since $\|\Phi(\theta,t)x\| \le Me^{2\omega}\|x\|$, for all $t \in [0,2)$, setting $K := \max\{(M^2e^{2\omega}L|\alpha|_I)/F_O(1), Me^{2\omega}\}$ we deduce that $\|\Phi(\theta,t)x\| \le K\|x\|$, for all $t \ge 0$. Taking into account that K does not depend on θ or x, it follows that

$$\|\Phi(\theta,t)x\| \le K\|x\|, \quad \forall t \ge 0,\ \forall x \in \mathcal{S}(\theta),\ \forall \theta \in \Theta. \tag{3.16}$$

(ii) Let $\theta \in \Theta$ and let $(x_n) \subset \mathcal{S}(\theta)$ with $x_n \underset{n\to\infty}{\longrightarrow} x$. For every $n \in \mathbb{N}$, we consider the sequence

$$\begin{gathered} v_n : \mathbb{R} \longrightarrow X, \quad v_n(t) = \alpha(t)\Phi(\theta,t)x_n, \\ f_n : \mathbb{R} \longrightarrow X, \quad f_n(t) = \begin{cases} \Phi(\theta,t)x_n, & t \ge 1, \\ \displaystyle\int_0^t \alpha(\tau)d\tau\,\Phi(\theta,t)x_n, & t \in [0,1), \\ 0, & t < 0. \end{cases} \end{gathered} \tag{3.17}$$

We have that $v_n \in C_{0c}(\mathbb{R}, X)$, for all $n \in \mathbb{N}$ and using similar arguments with those used in relation (3.10), we obtain that $f_n \in O(\mathbb{R}, X)$, for all $n \in \mathbb{N}$. An easy computation shows that the pair (f_n, v_n) satisfies (E_θ). Let $v : \mathbb{R} \to X$, $v(t) = \alpha(t)\Phi(\theta,t)x$. Then, $v \in C_{0c}(\mathbb{R}, X)$. According to our hypothesis there is, $f \in O(\mathbb{R}, X)$ such that the pair (f, v) satisfies (E_θ).

to Contents, page 1

Taking $u_n = v_n - v$ and $g_n = f_n - f$ we observe that $u_n \in \mathcal{C}_{0c}(\mathbb{R}, X)$, $g_n \in O(\mathbb{R}, X)$, and the pair (g_n, u_n) satisfies (E_θ). This implies that

$$\|f_n - f\|_{O(\mathbb{R},X)} \leq L\|v_n - v\|_{I(\mathbb{R},X)}, \quad \forall n \in \mathbb{N}. \tag{3.18}$$

From $\|v_n(t) - v(t)\| \leq \alpha(t)Me^{\omega}\|x_n - x\|$, for all $t \in \mathbb{R}$ and all $n \in \mathbb{N}$, we deduce that

$$\|v_n - v\|_{I(\mathbb{R},X)} \leq Me^{\omega}|\alpha|_I\|x_n - x\|, \quad \forall n \in \mathbb{N}. \tag{3.19}$$

From (3.18) and (3.19), it follows that $f_n \underset{n\to\infty}{\longrightarrow} f$ in $O(\mathbb{R}, X)$. From Remark 2.4(ii), we have that there is a subsequence (f_{k_n}) and a negligible set $A \subset \mathbb{R}$ such that $f_{k_n}(t) \underset{n\to\infty}{\longrightarrow} f(t)$, for all $t \in \mathbb{R} \setminus A$. In particular, it follows that there is $r > 1$ such that

$$f(r) = \lim_{n\to\infty} f_{k_n}(r) = \lim_{n\to\infty} \Phi(\theta, r)x_{k_n} = \Phi(\theta, r)x. \tag{3.20}$$

Because the pair (f, v) satisfies (E_θ), we obtain that

$$f(t) = \Phi(\sigma(\theta, r), t - r)f(r) = \Phi(\theta, t)x, \quad \forall t \geq r. \tag{3.21}$$

This shows that $f(t) = \lambda_{x,\theta}(t)$, for all $t \geq r$. Then, from

$$\|\lambda_{x,\theta}(t)\| \leq \|f(t)\| + Me^{\omega r}\|x\|\chi_{[0,r)}(t), \quad \forall t \in \mathbb{R}, \tag{3.22}$$

using the fact that $f \in O(\mathbb{R}, X)$ and Remark 2.4(i), we obtain that $\lambda_{x,\theta} \in O(\mathbb{R}, X)$, so $x \in \mathcal{S}(\theta)$.

In conclusion, $\mathcal{S}(\theta)$ is a closed linear subspace, for all $\theta \in \Theta$. □

Theorem 3.12 (The behavior on the unstable subspace). *If the pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ is uniformly admissible for the system (E_π), then the following assertions hold:*

(i) *there is $K > 0$ such that $\|\Phi(\theta, t)y\| \geq (1/K)\|y\|$, for all $t \geq 0$, all $y \in \mathcal{U}(\theta)$ and all $\theta \in \Theta$;*

(ii) *$\mathcal{U}(\theta)$ is a closed linear subspace, for all $\theta \in \Theta$.*

Proof. Let $L > 0$ be given by Definition 3.6 and let $M, \omega > 0$ be given by Definition 3.2. Let $\alpha : \mathbb{R} \to [0, 2]$ be a continuous function with $\operatorname{supp} \alpha \subset (0, 1)$ and $\int_0^1 \alpha(\tau)d\tau = 1$.

(i) Let $\theta \in \Theta$ and let $y \in \mathcal{U}(\theta)$. Then, there is $\varphi \in \mathcal{F}(\theta) \cap O(\mathbb{R}, X)$ with $\varphi(0) = y$. Let $t > 0$. We consider the functions

$$v : \mathbb{R} \longrightarrow X, \quad v(s) = -\alpha(s - t)\Phi(\theta, s)y,$$
$$f : \mathbb{R} \longrightarrow X, \quad f(s) = \begin{cases} \displaystyle\int_s^\infty \alpha(\tau - t)d\tau\, \Phi(\theta, s)y, & s \geq 0, \\ \varphi(s), & s < 0. \end{cases} \tag{3.23}$$

to Contents, page 1

We have that $v \in \mathcal{C}_{0c}(\mathbb{R}, X)$ and f is continuous. Let $m = \sup_{s\in[0,t+1]}\|f(s)\|$. Then, we have that

$$\|f(s)\| \leq \|\varphi(s)\| + m\chi_{[0,t+1]}(s), \quad \forall s \in \mathbb{R}. \tag{3.24}$$

From (3.24) and Remark 2.4(i), we deduce that $f \in O(\mathbb{R}, X)$. An easy computation shows that the pair (f, v) satisfies (E_θ). Then, according to our hypothesis, we have that

$$\|f\|_{O(\mathbb{R},X)} \leq L\|v\|_{I(\mathbb{R},X)}. \tag{3.25}$$

From $\|v(s)\| \leq \alpha(s-t)Me^{\omega}\|\Phi(\theta,t)y\|$, for all $s \in \mathbb{R}$, we obtain that

$$\|v\|_{I(\mathbb{R},X)} \leq |\alpha|_I Me^{\omega}\|\Phi(\theta,t)y\|. \tag{3.26}$$

Since $y = \varphi(0) = \Phi(\sigma(\theta,s),-s)\varphi(s)$, for all $s \in [-1,0)$, we have that

$$\|y\|\chi_{[-1,0)}(s) \leq Me^{\omega}\|\varphi(s)\|\chi_{[-1,0)}(s) \leq Me^{\omega}\|f(s)\|, \quad \forall s \in \mathbb{R}. \tag{3.27}$$

Using the invariance under translations of the space O from relation (3.27), we obtain that

$$\|y\|F_O(1) \leq Me^{\omega}\|f\|_{O(\mathbb{R},X)}. \tag{3.28}$$

Taking $K = (M^2e^{2\omega}L|\alpha|_I)/F_O(1)$ from relations (3.25), (3.26), and (3.28), it follows that $\|\Phi(\theta,t)y\| \geq (1/K)\|y\|$. Taking into account that K does not depend on t, y or θ, we conclude that

$$\|\Phi(\theta,t)y\| \geq \frac{1}{K}\|y\|, \quad \forall t \geq 0,\ \forall y \in \mathcal{U}(\theta),\ \forall \theta \in \Theta. \tag{3.29}$$

(ii) Let $\theta \in \Theta$ and let $(y_n) \subset \mathcal{U}(\theta)$ with $y_n \to y$. Then, for every $n \in \mathbb{N}$, there is $\varphi_n \in O(\mathbb{R}, X) \cap \mathcal{F}(\theta)$ with $\varphi_n(0) = y_n$. For every $n \in \mathbb{N}$, we consider the functions

$$\begin{aligned} &v_n : \mathbb{R} \longrightarrow X, \quad v_n(t) = -\alpha(t)\Phi(\theta,t)y_n, \\ &f_n : \mathbb{R} \longrightarrow X, \quad f_n(t) = \begin{cases} \displaystyle\int_t^{\infty} \alpha(\tau)d\tau\, \Phi(\theta,t)y_n, & t \geq 0, \\ \varphi_n(t), & t < 0. \end{cases} \end{aligned} \tag{3.30}$$

We have that $v_n \in \mathcal{C}_{0c}(\mathbb{R}, X)$, and, using similar arguments with those used in relation (3.24), we deduce that $f_n \in O(\mathbb{R}, X)$, for all $n \in \mathbb{N}$. An easy computation shows that the pair (f_n, v_n) satisfies (E_θ). Let

$$v : \mathbb{R} \longrightarrow X, \quad v(t) = -\alpha(t)\Phi(\theta,t)y. \tag{3.31}$$

to Contents, page 1

According to our hypothesis, there is $f \in O(\mathbb{R}, X)$ such that the pair (f, v) satisfies (E_θ). In particular, this implies that $f \in \mathcal{F}(\theta)$. Moreover, for every $n \in \mathbb{N}$, the pair $(f_n - f, v_n - v)$ satisfies (E_θ). According to our hypothesis, it follows that

$$\|f_n - f\|_{O(\mathbb{R},X)} \leq L\|v_n - v\|_{I(\mathbb{R},X)}, \quad \forall n \in \mathbb{N}. \tag{3.32}$$

We have that $\|v_n(t) - v(t)\| \leq \alpha(t) M e^{\omega}\|y_n - y\|$, for all $t \in \mathbb{R}$ and all $n \in \mathbb{N}$, so

$$\|v_n - v\|_{I(\mathbb{R},X)} \leq M e^{\omega}|\alpha|_I \|y_n - y\|, \quad \forall n \in \mathbb{N}. \tag{3.33}$$

From (3.32) and (3.33) it follows that $f_n \underset{n\to\infty}{\longrightarrow} f$ in $O(\mathbb{R}, X)$. Then, from Remark 2.4(ii), there is a subsequence $(f_{k_n}) \subset (f_n)$ and a negligible set $A \subset \mathbb{R}$ such that $f_{k_n}(t) \underset{n\to\infty}{\longrightarrow} f(t)$, for all $t \in \mathbb{R} \setminus A$. In particular, there is $h < 0$ such that $f_{k_n}(h) \underset{n\to\infty}{\longrightarrow} f(h)$. Since $f, f_{k_n} \in \mathcal{F}(\theta)$, we successively deduce that

$$y = \lim_{n\to\infty} y_{k_n} = \lim_{n\to\infty} f_{k_n}(0) = \lim_{n\to\infty} \Phi(\sigma(\theta, h), -h) f_{k_n}(h) = \Phi(\sigma(\theta, h), -h) f(h) = f(0). \tag{3.34}$$

This implies that $y \in \mathcal{U}(\theta)$, so $\mathcal{U}(\theta)$ is a closed linear subspace. □

Taking into account the above results it makes sense to study whether the uniform admissibility of a pair of function spaces from the class $\mathcal{T}(\mathbb{R})$ is a sufficient condition for the existence of the uniform dichotomy. Thus, the main result of this section is as follows.

Theorem 3.13 (Sufficient condition for uniform dichotomy). *Let $O, I \in \mathcal{T}(\mathbb{R})$ and let $\pi = (\Phi, \sigma)$ be a skew-product flow on $X \times \Theta$. If the pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ is uniformly admissible for the system (E_π), then π is uniformly dichotomic.*

Proof. Let $L > 0$ be given by Definition 3.6. Let $M, \omega > 0$ be given by Definition 3.2. Let $\alpha : \mathbb{R} \to [0, 2]$ be a continuous function with $\operatorname{supp} \alpha \subset (0, 1)$ and $\int_0^1 \alpha(\tau) d\tau = 1$. □

Step 1. We prove that $\mathcal{S}(\theta) \cap \mathcal{U}(\theta) = \{0\}$, for all $\theta \in \Theta$.

Let $\theta \in \Theta$ and let $x \in \mathcal{S}(\theta) \cap \mathcal{U}(\theta)$. Then, there is $\varphi \in O(\mathbb{R}, X) \cap \mathcal{F}(\theta)$ with $\varphi(0) = x$. We consider the function

$$f : \mathbb{R} \longrightarrow X, \qquad f(t) = \begin{cases} \Phi(\theta, t)x, & t \geq 0, \\ \varphi(t), & t < 0. \end{cases} \tag{3.35}$$

Then, $\|f(t)\| \leq \|\varphi(t)\| + \|\lambda_{x,\theta}(t)\|$, for all $t \in \mathbb{R}$. This implies that $f \in O(\mathbb{R}, X)$. An easy computation shows that the pair $(f, 0)$ satisfies (E_θ). Then, according to our hypothesis, it follows that $\|f\|_{O(\mathbb{R},X)} = 0$, so $f(t) = 0$ a.e. $t \in \mathbb{R}$. Observing that f is continuous, we obtain that $f(t) = 0$, for all $t \in \mathbb{R}$. In particular, we have that $x = f(0) = 0$.

Step 2. We prove that $\mathcal{S}(\theta) + \mathcal{U}(\theta) = X$, for all $\theta \in \Theta$.

to Contents, page 1

Let $\theta \in \Theta$ and let $x \in X$. Let $v : \mathbb{R} \to X$, $v(t) = \alpha(t)\Phi(\theta, t)x$. Then, $v \in \mathcal{C}_{0c}(\mathbb{R}, X)$, so there is $f \in O(\mathbb{R}, X)$ such that the pair (f, v) satisfies (E_θ). In particular, this implies that $f \in \mathcal{F}(\theta)$, so $f(0) \in \mathcal{U}(\theta)$. In addition, we observe that

$$f(t) = \Phi(\theta, t)f(0) + \left(\int_0^t \alpha(\tau)d\tau\right)\Phi(\theta, t)x = \Phi(\theta, t)\big(f(0) + x\big), \quad \forall t \geq 1. \tag{3.36}$$

Setting $z_x = f(0) + x$ from (3.36), we have that $\lambda_{z_x,\theta}(t) = f(t)$, for all $t \geq 1$. It follows that

$$\|\lambda_{z_x,\theta}(t)\| \leq \|f(t)\| + Me^{\omega}\|z_x\|\chi_{[0,1)}(t), \quad \forall t \in \mathbb{R}. \tag{3.37}$$

From relation (3.37) and Remark 2.4(i) we obtain that $\lambda_{z_x,\theta} \in O(\mathbb{R}, X)$, so $z_x \in \mathcal{S}(\theta)$. This shows that $x = z_x - f(0) \in \mathcal{S}(\theta) + \mathcal{U}(\theta)$, so $\mathcal{S}(\theta) + \mathcal{U}(\theta) = X$.

According to Steps 1 and 2, Theorem 3.11(ii), and Theorem 3.12(ii), we deduce that

$$\mathcal{S}(\theta) \oplus \mathcal{U}(\theta) = X, \quad \forall \theta \in \Theta. \tag{3.38}$$

For every $\theta \in \Theta$ we denote by $P(\theta)$ the projection with the property that

$$\operatorname{Range} P(\theta) = \mathcal{S}(\theta), \qquad \operatorname{Ker} P(\theta) = \mathcal{U}(\theta). \tag{3.39}$$

Using Proposition 3.9 we obtain that

$$\Phi(\theta, t)P(\theta) = P(\sigma(\theta, t))\Phi(\theta, t), \quad \forall(\theta, t) \in \Theta \times \mathbb{R}_+. \tag{3.40}$$

Let $(\theta, t) \in \Theta \times \mathbb{R}_+$. From Proposition 3.9(ii), it follows that the restriction $\Phi(\theta, t)_| : \operatorname{Ker} P(\theta) \to \operatorname{Ker} P(\sigma(\theta, t))$ is correctly defined and surjective. According to Theorem 3.12(ii) we have that $\Phi(\theta, t)_|$ is also injective, so this is an isomorphism, for all $(\theta, t) \in \Theta \times \mathbb{R}_+$.

Step 3. We prove that $\sup_{\theta\in\Theta}\|P(\theta)\| < \infty$.

Let $\theta \in \Theta$ and let $x \in X$. Let $x_s^\theta = P(\theta)x$ and let $x_u^\theta = (I - P(\theta))x$. Since $x_u^\theta \in \operatorname{Ker} P(\theta) = \mathcal{U}(\theta)$, there is $\varphi \in \mathcal{F}(\theta) \cap O(\mathbb{R}, X)$ with $\varphi(0) = x_u^\theta$. We consider the functions

$$v : \mathbb{R} \longrightarrow X, \quad v(t) = \alpha(t)\Phi(\theta, t)x,$$
$$f : \mathbb{R} \to X, \quad f(t) = \begin{cases} \Phi(\theta, t)x_s^\theta, & t \geq 1, \\ -\Phi(\theta, t)x_u^\theta + \left(\int_0^t \alpha(\tau)d\tau\right)\Phi(\theta, t)x, & t \in [0, 1), \\ -\varphi(t), & t < 0. \end{cases} \tag{3.41}$$

to Contents, page 1

We have that $v \in C_{0c}(\mathbb{R}, X)$ and f is continuous. From $x_s^\theta \in \operatorname{Range} P(\theta) = \mathcal{S}(\theta)$, we have that the function $\lambda_{x_s^\theta,\theta}$ belongs to $O(\mathbb{R}, X)$. Setting $m = \sup_{t\in[0,1]}\|f(t)\|$ and observing that

$$\|f(t)\| \leq \|\varphi(t)\| + m\chi_{[0,1]}(t) + \|\lambda_{x_s^\theta,\theta}(t)\|, \quad \forall t \in \mathbb{R}, \tag{3.42}$$

from (3.42), we deduce that $f \in O(\mathbb{R}, X)$. An easy computation shows that the pair (f, v) satisfies (E_θ). This implies that

$$\|f\|_{O(\mathbb{R},X)} \leq L\|v\|_{I(\mathbb{R},X)}. \tag{3.43}$$

Since $\varphi \in \mathcal{F}(\theta)$, we have that $x_u^\theta = \varphi(0) = \Phi(\sigma(\theta, s), -s)\varphi(s)$, for all $s \in [-1, 0)$. This implies that

$$\|x_u^\theta\| \leq Me^\omega\|\varphi(s)\| = Me^\omega\|f(s)\|, \quad \forall s \in [-1, 0), \tag{3.44}$$

and we obtain that

$$\|x_u^\theta\|\chi_{[-1,0)}(s) \leq Me^\omega\|f(s)\|, \quad \forall s \in \mathbb{R}. \tag{3.45}$$

Using the invariance under translations of the space O, from relation (3.45) we deduce that

$$\|x_u^\theta\|F_O(1) \leq Me^\omega\|f\|_{O(\mathbb{R},X)}. \tag{3.46}$$

In addition, from

$$\|v(t)\| \leq \alpha(t)Me^\omega\|x\|, \quad \forall t \in \mathbb{R}, \tag{3.47}$$

we obtain that

$$\|v\|_{I(\mathbb{R},X)} \leq |\alpha|_I Me^\omega\|x\|. \tag{3.48}$$

Setting $\gamma := [L|\alpha|_I M^2e^{2\omega}/F_O(1)]$ from relations (3.43), (3.46), and (3.48), we have that

$$\|(I - P(\theta))x\| = \|x_u^\theta\| \leq \gamma\|x\|. \tag{3.49}$$

This implies that

$$\|P(\theta)x\| \leq (1 + \gamma)\|x\|. \tag{3.50}$$

Taking into account that γ does not depend on θ or x, it follows that relation (3.50) holds, for all $\theta \in \Theta$ and all $x \in X$, so $\|P(\theta)\| \leq 1 + \gamma$, for all $\theta \in \Theta$.

to Contents, page 1

Finally, from Theorem 3.11(i) and Theorem 3.12(i), we conclude that π is uniformly dichotomic.

Remark 3.14. Relation (3.39) shows that the stable subspace and the instable subspace play a central role in the detection of the dichotomous behavior of a skew-product flow and gives a comprehensible motivation for their usual appellation.

4. Exponential Dichotomy of Skew-Product Flows

In the previous section, we have obtained sufficient conditions for the uniform dichotomy of a skew-product flow $\pi = (\Phi, \sigma)$ on $X \times \Theta$ in terms of the uniform admissibility of the pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ for the associated control system (E_π), where $O, I \in \mathcal{T}(\mathbb{R})$. The natural question arises: which are the additional (preferably minimal) hypotheses under which this admissibility may provide the existence of the *exponential* dichotomy? In this context, the main purpose of this section is to establish which are the most general classes of Banach function spaces where O or I may belong to, such that the uniform admissibility of the pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ for the control system (E_π) is a sufficient (and also a necessary) condition for the existence of exponential dichotomy.

Let X be a real or complex Banach space and let (Θ, d) be a metric space. Let $\pi = (\Phi, \sigma)$ be a skew-product flow on $X \times \Theta$.

Definition 4.1. A skew-product flow $\pi = (\Phi, \sigma)$ is said to be *exponentially dichotomic* if there exist a family of projections $\{P(\theta)\}_{\theta\in\Theta} \subset \mathcal{B}(X)$ and two constants $K \geq 1$ and $\nu > 0$ such that the following properties hold:

(i) $\Phi(\theta, t)P(\theta) = P(\sigma(\theta, t))\Phi(\theta, t)$, for all $(\theta, t) \in \Theta \times \mathbb{R}_+$;

(ii) $\|\Phi(\theta, t)x\| \leq Ke^{-\nu t}\|x\|$, for all $t \geq 0$, all $x \in \operatorname{Range} P(\theta)$ and all $\theta \in \Theta$;

(iii) the restriction $\Phi(\theta, t)_| : \operatorname{Ker} P(\theta) \to \operatorname{Ker} P(\sigma(\theta, t))$ is an isomorphism, for all $(\theta, t) \in \Theta \times \mathbb{R}_+$;

(iv) $\|\Phi(\theta, t)y\| \geq (1/K)e^{\nu t}\|y\|$, for all $t \geq 0$, all $y \in \operatorname{Ker} P(\theta)$ and all $\theta \in \Theta$.

Before proceeding to the next steps, we need a technical lemma.

Lemma 4.2. *If a skew-product flow π is exponentially dichotomic with respect to a family of projections $\{P(\theta)\}_{\theta\in\Theta}$, then $\sup_{\theta\in\Theta}\|P(\theta)\| < \infty$.*

Proof. Let $K, \nu > 0$ be given by Definition 4.1 and let $M, \omega > 0$ be given by Definition 3.2. For every $(x, \theta) \in X \times \Theta$ and every $t \geq 0$, we have that

$$\begin{aligned} \frac{1}{K}e^{\nu t}\|(I - P(\theta))x\| &\leq \|\Phi(\theta, t)(I - P(\theta))x\| \leq Me^{\omega t}\|x\| + Ke^{-\nu t}\|P(\theta)x\| \\ &\leq \left(Me^{\omega t} + K\right)\|x\| + Ke^{-\nu t}\|(I - P(\theta))x\|, \end{aligned} \tag{4.1}$$

which implies that

$$\left(e^{2\nu t} - K^2\right)\frac{e^{-\nu t}}{K}\|(I - P(\theta))x\| \leq \left(Me^{\omega t} + K\right)\|x\|, \quad \forall t \geq 0,\ \forall (x, \theta) \in X \times \Theta. \tag{4.2}$$

to Contents, page 1

Let $h > 0$ be such that $e^{2\nu h} - K^2 > 0$. Setting $\alpha := (e^{2\nu h} - K^2)e^{-\nu h}/K$ and $\delta := (Me^{\omega h} + K)$, it follows that $\|(I-P(\theta))x\| \le (\delta/\alpha)\|x\|$, for all $(x,\theta) \in X\times\Theta$. This implies that $\|I-P(\theta)\| \le \delta/\alpha$, for all $\theta \in \Theta$, so $\|P(\theta)\| \le 1 + (\delta/\alpha)$, for all $\theta \in \Theta$, and the proof is complete. □

Remark 4.3. (i) Using Lemma 4.2, we deduce that if a skew-product flow π is exponentially dichotomic with respect to a family of projections $\{P(\theta)\}_{\theta\in\Theta}$, then π is uniformly dichotomic with respect to the same family of projections.

(ii) If a skew-product flow π is exponentially dichotomic with respect to a family of projections $\{P(\theta)\}_{\theta\in\Theta}$, then this family is uniquely determined (see, e.g., [18], Remark 2.5).

Remark 4.4. In the description of any dichotomous behavior, the properties (i) and (iii) are inherent, because beside the splitting of the space ensured by the presence of the dichotomy projections, these properties reflect both the invariance with respect to the decomposition induced by each projection as well as the reversibility of the cocycle restricted to the kernel of each projection.

In this context, it is extremely important to note that if in the detection of the dichotomy one assumes from the very beginning that there exist a projection family such that the invariance property (i) and the reversibility condition (iii) hold, then the dichotomy concept is resumed to a stability property (ii) and to an instability condition (iv), which via (iii) will consist only of a double stability. Thus, if in the study of the dichotomy one considers (i) and (iii) as working hypotheses, then the entire investigation is reduced to a quasitrivial case of (double) stability.

In conclusion, in the study of the existence of (uniform or) exponential dichotomy, it is essential to determine conditions *which imply the existence of the projection family* and also the fulfillment of *all* the conditions from Definition 4.1.

Now let O, I be two Banach function spaces such that $O, I \in \mathcal{T}(\mathbb{R})$. According to the main result in the previous section (see Theorem 3.13), if the pair $(O(\mathbb{R},X), I(\mathbb{R},X))$ is uniformly admissible for the system (E_π), then π is uniformly dichotomic with respect to a family of projections $\{P(\theta)\}_{\theta\in\Theta}$ with the property that

$$\operatorname{Range} P(\theta) = \mathcal{S}(\theta), \qquad \operatorname{Ker} P(\theta) = \mathcal{U}(\theta), \quad \forall\theta \in \Theta. \tag{4.3}$$

In what follows, we will see that by imposing some conditions *either* on the output space O *or* on the input space I, the admissibility becomes a sufficient condition for the exponential dichotomy.

Theorem 4.5 (The behavior on the stable subspace). *Let O, I be two Banach function spaces such that either $O \in \mathcal{Q}(\mathbb{R})$ or $I \in \mathcal{L}(\mathbb{R})$. If the pair $(O(\mathbb{R},X), I(\mathbb{R},X))$ is uniformly admissible for the system (E_π), then there are $K, \nu > 0$ such that*

$$\|\Phi(\theta,t)x\| \le Ke^{-\nu t}\|x\|, \quad \forall t \ge 0,\ \forall x \in \operatorname{Range} P(\theta),\ \forall\theta \in \Theta. \tag{4.4}$$

Proof. Let $\delta > 0$ be such that

$$\|\Phi(\theta,t)x\| \le \delta\|x\|, \quad \forall t \ge 0,\ \forall x \in \operatorname{Range} P(\theta),\ \forall\theta \in \Theta. \tag{4.5}$$

to Contents, page 1

We prove that there is $h > 0$ such that

$$\|\Phi(\theta, h)x\| \le \frac{1}{e}\|x\|, \quad \forall x \in \operatorname{Range} P(\theta),\ \forall\theta \in \Theta. \tag{4.6}$$

Let $L > 0$ be given by Definition 3.6 and let $M, \omega > 0$ be given by Definition 3.2.

Case 1. Suppose that $O \in \mathcal{Q}(\mathbb{R})$. Let $\alpha : \mathbb{R} \to [0,2]$ be a continuous function with $\operatorname{supp}\alpha \subset (0,1)$ such that $\int_0^1 \alpha(\tau)d\tau = 1$. Since $\sup_{t>0} F_O(t) = \infty$, there is $r > 0$ such that

$$F_O(r) \ge e\delta^2 L|\alpha|_I. \tag{4.7}$$

Let $\theta \in \Theta$ and let $x \in \operatorname{Range} P(\theta)$. If $\Phi(\theta,1)x \neq 0$, then we consider the functions

$$\begin{gathered} v : \mathbb{R} \longrightarrow X, \quad v(t) = \alpha(t)\frac{\Phi(\theta,t)x}{\|\Phi(\theta,t)x\|}, \\ f : \mathbb{R} \longrightarrow X, \quad f(t) = \begin{cases} a\Phi(\theta,t)x, & t \ge 1, \\ \displaystyle\int_0^t \frac{\alpha(\tau)}{\|\Phi(\theta,\tau)x\|}d\tau\, \Phi(\theta,t)x, & t \in [0,1], \\ 0, & t < 0, \end{cases} \end{gathered} \tag{4.8}$$

where

$$a := \int_0^1 \frac{\alpha(\tau)}{\|\Phi(\theta,\tau)x\|}d\tau. \tag{4.9}$$

We observe that f is continuous and

$$\|f(t)\| \le a\|\lambda_{x,\theta}(t)\|, \quad \forall t \in \mathbb{R}. \tag{4.10}$$

Since $x \in \operatorname{Range} P(\theta) = \mathcal{S}(\theta)$, we have that $\lambda_{x,\theta} \in O(\mathbb{R}, X)$. Then using Remark 2.4(i), we deduce that $f \in O(\mathbb{R}, X)$. In addition, we have that $v \in C_{0c}(\mathbb{R}, X)$ and an easy computation shows that the pair (f, v) satisfies (E_θ). Then, according to our hypothesis, it follows that

$$\|f\|_{O(\mathbb{R},X)} \le L\|v\|_{I(\mathbb{R},X)}. \tag{4.11}$$

Because $\|v(t)\| = \alpha(t)$, for all $t \in \mathbb{R}$, the relation (4.11) becomes

$$\|f\|_{O(\mathbb{R},X)} \le L|\alpha|_I. \tag{4.12}$$

Using relation (4.5), we deduce that

$$\|\Phi(\theta, r+1)x\| \le \delta\|\Phi(\theta,t)x\| = \frac{\delta}{a}\|f(t)\|, \quad \forall t \in [1, r+1), \tag{4.13}$$

to Contents, page 1

so

$$\|\Phi(\theta, r+1)x\|\chi_{[1,r+1)}(t) \leq \frac{\delta}{a}\|f(t)\|, \quad \forall t \in \mathbb{R}. \tag{4.14}$$

Using the invariance under translations of the space O from relation (4.14), we obtain that

$$\|\Phi(\theta, r+1)x\|F_O(r) \leq \frac{\delta}{a}\|f\|_{O(\mathbb{R},X)}. \tag{4.15}$$

Setting $h := r+1$ from relations (4.12) and (4.15), it follows that

$$\|\Phi(\theta, h)x\|F_O(r) \leq \frac{\delta L|\alpha|_I}{a}. \tag{4.16}$$

Moreover, from relation (4.5), we have that $\|\Phi(\theta,\tau)x\| \leq \delta\|x\|$, for all $\tau \in [0,1)$, so

$$a = \int_0^1 \frac{\alpha(\tau)}{\|\Phi(\theta,\tau)x\|}d\tau \geq \frac{1}{\delta\|x\|}. \tag{4.17}$$

From relations (4.7), (4.16), and (4.17), it follows that

$$\|\Phi(\theta, h)x\| \leq \frac{1}{e}\|x\|. \tag{4.18}$$

If $\Phi(\theta,1)x = 0$, then $\Phi(\theta,h)x = 0$, so the above relation holds. Taking into account that h does not depend on θ or x, we obtain that in this case, there is $h > 0$ such that relation (4.6) holds.

Case 2. Suppose that $I \in \mathcal{L}(\mathbb{R})$. In this situation, from Remark 2.16, we have that there is a continuous function $\gamma : \mathbb{R} \to \mathbb{R}_+$ such that $\gamma \in I \setminus L^1(\mathbb{R},\mathbb{R})$. Since the space I is invariant under translations, we may assume that there is $r > 1$ such that

$$\int_1^r \gamma(\tau)d\tau \geq \frac{eL\delta^2|\gamma|_I}{F_O(1)}. \tag{4.19}$$

Let $\beta : \mathbb{R} \to [0,1]$ be a continuous function with $\operatorname{supp}\beta \subset (0, r+1)$ and $\beta(t) = 1$, for all $t \in [1,r]$.

Let $\theta \in \Theta$ and let $x \in \operatorname{Range} P(\theta)$. We consider the functions

$$\begin{aligned}
&v : \mathbb{R} \longrightarrow X, \quad v(t) = \beta(t)\gamma(t)\Phi(\theta,t)x,\\
&f : \mathbb{R} \longrightarrow X, \quad f(t) = \begin{cases} q\Phi(\theta,t)x, & t \geq r+1,\\ \displaystyle\int_0^t \beta(\tau)\gamma(\tau)d\tau\,\Phi(\theta,t)x, & t \in [0,r+1),\\ 0, & t < 0,\end{cases}
\end{aligned} \tag{4.20}$$

to Contents, page 1

where

$$q = \int_0^{r+1} \beta(\tau)\gamma(\tau)d\tau. \tag{4.21}$$

We have that $v \in C_{0c}(\mathbb{R}, X)$, f is continuous, and $\|f(t)\| \le q\|\lambda_{x,\theta}(t)\|$, for all $t \in \mathbb{R}$. Using similar arguments with those used in relation (4.10), we deduce that $f \in O(\mathbb{R}, X)$. An easy computation shows that the pair (f, v) satisfies (E_θ). Then, we have that

$$\|f\|_{O(\mathbb{R},X)} \le L\|v\|_{I(\mathbb{R},X)}. \tag{4.22}$$

Using relation (4.5), we obtain that

$$\|v(t)\| \le \delta\gamma(t)\|x\|, \quad \forall t \in \mathbb{R}, \tag{4.23}$$

which implies that

$$\|v\|_{I(\mathbb{R},X)} \le \delta|\gamma|_I\|x\|. \tag{4.24}$$

In addition, from $\|\Phi(\theta, r+2)x\| \le \delta\|\Phi(\theta, t)x\|$, for all $t \in [r+1, r+2)$, we deduce that

$$\|\Phi(\theta, r+2)x\|\chi_{[r+1,r+2)}(t) \le \frac{\delta}{q}\|f(t)\|, \quad \forall t \in \mathbb{R}. \tag{4.25}$$

Using the invariance under translations of the space O from relations (4.25), (4.22), and (4.24) we have that

$$q\|\Phi(\theta, r+2)x\|F_O(1) \le \delta\|f\|_{O(\mathbb{R},X)} \le L\delta^2|\gamma|_I\|x\|. \tag{4.26}$$

Since $q \ge \int_1^r \gamma(\tau)d\tau$, from relations (4.19), (4.21), and (4.26), it follows that

$$\|\Phi(\theta, r+2)x\| \le \frac{1}{e}\|x\|. \tag{4.27}$$

Setting $h = r+2$ and taking into account that h does not depend on θ or x, we obtain that relation (4.6) holds.

In conclusion, in both situations, there is $h > 0$ such that

$$\|\Phi(\theta, h)x\| \le \frac{1}{e}\|x\|, \quad \forall x \in \text{Range}\, P(\theta),\ \forall\theta \in \Theta. \tag{4.28}$$

to Contents, page 1

Let $\nu := 1/h$ and let $K = \delta e$. Let $\theta \in \Theta$ and let $x \in \operatorname{Range} P(\theta)$. Let $t > 0$. Then, there are $k \in \mathbb{N}$ and $\tau \in [0, h)$ such that $t = kh + \tau$. Using relations (4.5) and (4.6), we successively deduce that

$$\|\Phi(\theta, t)x\| \leq \delta\|\Phi(\theta, kh)x\| \leq \delta e^{-k}\|x\| \leq K e^{-\nu t}\|x\|. \tag{4.29}$$

□

Theorem 4.6 (The behavior on the unstable subspace). *Let O, I be two Banach function spaces such that either $O \in \mathcal{Q}(\mathbb{R})$ or $I \in \mathcal{L}(\mathbb{R})$. If the pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ is uniformly admissible for the system (E_π), then, there are $K, \nu > 0$ such that*

$$\|\Phi(\theta, t)y\| \geq \frac{1}{K} e^{\nu t}\|y\|, \quad \forall t \geq 0,\ \forall y \in \operatorname{Ker} P(\theta),\ \forall \theta \in \Theta. \tag{4.30}$$

Proof. Let $\delta > 0$ be such that

$$\|\Phi(\theta, t)y\| \geq \frac{1}{\delta}\|y\|, \quad \forall t \geq 0,\ \forall y \in \operatorname{Ker} P(\theta),\ \forall \theta \in \Theta. \tag{4.31}$$

Let $L > 0$ be given by Definition 3.6 and let $M, \omega > 0$ be given by Definition 3.2. We prove that there is $h > 0$ such that

$$\|\Phi(\theta, h)y\| \geq e\|y\|, \quad \forall y \in \operatorname{Ker} P(\theta),\ \forall \theta \in \Theta. \tag{4.32}$$

Case 1. Suppose that $O \in \mathcal{Q}(\mathbb{R})$. Let $\alpha : \mathbb{R} \to [0, 2]$ be a continuous function with $\operatorname{supp}\alpha \subset (0, 1)$ and $\int_0^1 \alpha(\tau)d\tau = 1$. In this case, there is $r > 0$ such that

$$F_O(r) \geq e\delta^2 L|\alpha|_I. \tag{4.33}$$

Let $\theta \in \Theta$ and let $y \in \operatorname{Ker} P(\theta) \setminus \{0\}$. Then, $\Phi(\theta, t)y \neq 0$, for all $t \geq 0$. Since $y \in \operatorname{Ker} P(\theta) = \mathcal{U}(\theta)$, there is $\varphi \in \mathcal{F}(\theta) \cap O(\mathbb{R}, X)$ with $\varphi(0) = y$. We consider the functions

$$\begin{gathered} v : \mathbb{R} \longrightarrow X, \quad v(t) = -\alpha(t - r)\frac{\Phi(\theta, t)y}{\|\Phi(\theta, t)y\|} \\ f : \mathbb{R} \longrightarrow X, \quad f(t) = \begin{cases} \displaystyle\int_t^\infty \frac{\alpha(\tau - r)}{\|\Phi(\theta, \tau)y\|} d\tau\, \Phi(\theta, t)y, & t \geq r, \\ a\Phi(\theta, t)y, & t \in [0, r), \\ a\varphi(t), & t < 0, \end{cases} \end{gathered} \tag{4.34}$$

where

$$a := \int_r^{r+1} \frac{\alpha(\tau - r)}{\|\Phi(\theta, \tau)y\|} d\tau. \tag{4.35}$$

to Contents, page 1

We have that $v \in \mathcal{C}_{0c}(\mathbb{R}, X)$ and f is continuous. Moreover, from

$$\|f(t)\| \leq a\|\varphi(t)\| + aMe^{\omega(r+1)}\|y\|\chi_{[0,r+1)}(t), \quad \forall t \in \mathbb{R}, \tag{4.36}$$

we obtain that $f \in O(\mathbb{R}, X)$. An easy computation shows that the pair (f, v) satisfies (E_θ), so

$$\|f\|_{O(\mathbb{R},X)} \leq L\|v\|_{I(\mathbb{R},X)}. \tag{4.37}$$

Observing that $\|v(t)\| = \alpha(t - r)$, for all $t \in \mathbb{R}$, the relation (4.37) becomes

$$\|f\|_{O(\mathbb{R},X)} \leq L|\alpha|_I. \tag{4.38}$$

From relation (4.31), we have that

$$\|\Phi(\theta, r+1)y\| \geq \frac{1}{\delta}\|\Phi(\theta, \tau)y\|, \quad \forall \tau \in [r, r+1]. \tag{4.39}$$

This implies that

$$a \geq \frac{1}{\delta\|\Phi(\theta, r+1)y\|}. \tag{4.40}$$

In addition, from relation (4.31), we have that

$$\|\Phi(\theta, t)y\| \geq \frac{1}{\delta}\|y\|, \quad \forall t \in [0, r) \tag{4.41}$$

which implies that

$$\|y\|\chi_{[0,r)}(t) \leq \delta\|\Phi(\theta, t)y\|\chi_{[0,r)}(t) \leq \frac{\delta}{a}\|f(t)\|, \quad \forall t \in \mathbb{R}. \tag{4.42}$$

From relation (4.42), it follows that

$$\|y\|F_O(r) \leq \frac{\delta}{a}\|f\|_{O(\mathbb{R},X)}. \tag{4.43}$$

From relations (4.38), (4.40), and (4.43), we deduce that

$$\|y\|F_O(r) \leq \frac{\delta L|\alpha|_I}{a} \leq \delta^2 L|\alpha|_I\|\Phi(\theta, r+1)y\|. \tag{4.44}$$

From relations (4.44) and (4.33), we have that

$$\|\Phi(\theta, r+1)y\| \geq e\|y\|. \tag{4.45}$$

to Contents, page 1

Setting $h := r + 1$ and taking into account that h does not depend on y or θ we obtain that relation (4.32) holds.

Case 2. Suppose that $I \in \mathcal{L}(\mathbb{R})$. In this situation, using Remark 2.16 and the translation invariance of the space I, we have that there is a continuous function $\gamma : \mathbb{R} \to \mathbb{R}_+$ with $\gamma \in I \setminus L^1(\mathbb{R}, \mathbb{R})$ and $r > 1$ such that

$$\int_1^r \gamma(\tau)d\tau \geq e^{\omega+1}\frac{LM\delta|\gamma|_I}{F_O(1)}. \tag{4.46}$$

Let $\beta : \mathbb{R} \to [0,1]$ be a continuous function with $\operatorname{supp}\beta \subset (0, r+1)$ and $\beta(t) = 1$, for all $t \in [1, r]$.

Let $\theta \in \Theta$ and let $y \in \operatorname{Ker} P(\theta)$. Since $\operatorname{Ker} P(\theta) = \mathcal{U}(\theta)$ there is $\varphi \in \mathcal{F}(\theta) \cap O(\mathbb{R}, X)$ with $\varphi(0) = y$. We consider the functions

$$\begin{gathered} v : \mathbb{R} \longrightarrow X, \quad v(t) = -\beta(t)\gamma(t)\Phi(\theta, t)y, \\ f : \mathbb{R} \longrightarrow X, \quad f(t) = \begin{cases} \displaystyle\int_t^\infty \beta(\tau)\gamma(\tau)d\tau\Phi(\theta, t)y, & t \geq 0, \\ q\varphi(t), & t < 0, \end{cases} \end{gathered} \tag{4.47}$$

where $q := \int_0^{r+1} \beta(\tau)\gamma(\tau)d\tau$. We have that $v \in C_{0c}(\mathbb{R}, X)$, and, using similar arguments with those from Case 1, we obtain that $f \in O(\mathbb{R}, X)$. An easy computation shows that the pair (f, v) satisfies (E_θ), so

$$\|f\|_{O(\mathbb{R},X)} \leq L\|v\|_{I(\mathbb{R},X)}. \tag{4.48}$$

From (4.31), we have that $\|\Phi(\theta, r+1)y\| \geq (1/\delta)\|\Phi(\theta, t)y\|$, for all $t \in [0, r+1]$. This implies that

$$\|v(t)\| \leq \gamma(t)\delta\|\Phi(\theta, r+1)y\|, \quad \forall t \in \mathbb{R}, \tag{4.49}$$

so

$$\|v\|_{I(\mathbb{R},X)} \leq |\gamma|_I\delta\|\Phi(\theta, r+1)y\|. \tag{4.50}$$

Since $\varphi \in \mathcal{F}(\theta)$, we have that

$$\|y\| = \|\varphi(0)\| = \|\Phi(\sigma(\theta, t), -t)\varphi(t)\| \leq Me^\omega\|\varphi(t)\|, \quad \forall t \in [-1, 0). \tag{4.51}$$

From relation (4.51), it follows that

$$\|y\|\chi_{[-1,0)}(t) \leq Me^\omega\|\varphi(t)\|\chi_{[-1,0)}(t) \leq \frac{Me^\omega}{q}\|f(t)\|, \quad \forall t \in \mathbb{R}. \tag{4.52}$$

to Contents, page 1

Using the translation invariance of the space O from (4.52), we obtain that

$$q\|y\|F_O(1) \le Me^{\omega}\|f\|_{O(\mathbb{R},X)}. \tag{4.53}$$

Since $q \ge \int_1^r \gamma(\tau)d\tau$, from relations (4.46), (4.48), (4.50) we deduce that

$$\|\Phi(\theta, r+1)y\| \ge e\|y\|. \tag{4.54}$$

Setting $h := r+1$ and since h does not depend on y or θ, we have that the relation (4.32) holds.

In conclusion, in both situations there is $h > 0$ such that

$$\|\Phi(\theta, h)y\| \ge e\|y\|, \quad \forall y \in \operatorname{Ker} P(\theta),\ \forall \theta \in \Theta. \tag{4.55}$$

Let $\nu = 1/h$ and let $K = \delta e$. Let $\theta \in \Theta$ and let $y \in \operatorname{Ker} P(\theta)$. Let $t > 0$. Then, there are $j \in \mathbb{N}$ and $s \in [0, h)$ such that $t = jh + s$. Using relations (4.31) and (4.32), we obtain that

$$\|\Phi(\theta, t)y\| \ge \frac{1}{\delta}\|\Phi(\theta, jh)y\| \ge \frac{1}{\delta}e^j\|y\| \ge \frac{1}{K}e^{\nu t}\|y\|. \tag{4.56}$$

□

According to the previous results we may formulate now a sufficient condition for the existence of the exponential dichotomy. Moreover, for the converse implication we will show that it sufficient to chose one of the spaces in the admissible pair from the class $\mathcal{R}(\mathbb{R})$. Thus, the main result of this section is as follows.

Theorem 4.7 (Necessary and sufficient condition for exponential dichotomy). *Let $\pi = (\Phi, \sigma)$ be a skew-product flow on $\mathcal{E} = X \times \Theta$ and let O, I be two Banach function spaces with $O, I \in \mathcal{T}(\mathbb{R})$ such that either $O \in \mathcal{Q}(\mathbb{R})$ or $I \in \mathcal{L}(\mathbb{R})$. The following assertions hold:*

(i) if the pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ is uniformly admissible for the system (E_π), then π is exponentially dichotomic.

(ii) if $I \subset O$ and one of the spaces I or O belongs to the class $\mathcal{R}(\mathbb{R})$, then π is exponentially dichotomic if and only if the pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ is uniformly admissible for the system (E_π).

Proof. (i) This follows from Theorem 3.13, Theorem 4.5, and Theorem 4.6.

(ii) Since $I \subset O$, it follows that there is $\alpha > 0$ such that

$$|u|_O \le \alpha |u|_I, \quad \forall u \in I. \tag{4.57}$$

Necessity. Suppose that π is exponentially dichotomic with respect to the family of projections $\{P(\theta)\}_{\theta\in\Theta}$ and let $K, \nu > 0$ be two constants given by Definition 4.1. According to Lemma 4.2, we have that $q := \sup_{\theta\in\Theta}\|P(\theta)\| < \infty$. For every $(\theta, t) \in \Theta \times \mathbb{R}_+$ we denote by $\Phi(\theta, t)_|^{-1}$ the inverse of the operator $\Phi(\theta, t)_| : \operatorname{Ker} P(\theta) \to \operatorname{Ker} P(\sigma(\theta, t))$.

to Contents, page 1

Let $\theta \in \Theta$ and let $v \in C_{0c}(\mathbb{R}, X)$. We consider the function $f_v : \mathbb{R} \to X$ given by

$$f_v(t) = \int_{-\infty}^{t} \Phi(\sigma(\theta,\tau), t-\tau) P(\sigma(\theta,\tau)) v(\tau) d\tau \\ - \int_{t}^{\infty} \Phi(\sigma(\theta,t), \tau - t)_{|}^{-1} (I - P(\sigma(\theta,\tau)))\ v(\tau) d\tau. \tag{4.58}$$

We have that f_v is continuous, and a direct computation shows that the pair (f_v, v) satisfies (E_θ). In addition, we have that

$$\|f_v(t)\| \leq qK \int_{-\infty}^{t} e^{-\nu(t-\tau)} \|v(\tau)\| d\tau \\ + (1+q)K \int_{t}^{\infty} e^{-\nu(\tau-t)} \|v(\tau)\| d\tau, \quad \forall t \in \mathbb{R}. \tag{4.59}$$

If $I \in \mathcal{R}(\mathbb{R})$, let $\gamma_{I,\nu} > 0$ be the constant given by Lemma 2.21. Then, from (4.59) and Lemma 2.21, it follows that $f_v \in I(\mathbb{R}, X)$ and

$$\|f_v\|_{I(\mathbb{R},X)} \leq (1+2q) K \gamma_{I,\nu} \|v\|_{I(\mathbb{R},X)}. \tag{4.60}$$

Then, from (4.57) and (4.60), we deduce that $f_v \in O(\mathbb{R}, X)$ and

$$\|f_v\|_{O(\mathbb{R},X)} \leq \alpha(1+2q) K \gamma_{I,\nu} \|v\|_{I(\mathbb{R},X)}. \tag{4.61}$$

If $O \in \mathcal{R}(\mathbb{R})$, let $\gamma_{O,\nu} > 0$ be the constant given by Lemma 2.21. Then, from (4.59), (4.57) and using Lemma 2.21, we successively obtain that $f_v \in O(\mathbb{R}, X)$ and

$$\|f_v\|_{O(\mathbb{R},X)} \leq (1+2q) K \gamma_{O,\nu} \|v\|_{O(\mathbb{R},X)} \leq \alpha(1+2q) K \gamma_{O,\nu} \|v\|_{I(\mathbb{R},X)}. \tag{4.62}$$

Let

$$\gamma := \begin{cases} \gamma_{I,\nu}, & \text{if } I \in \mathcal{R}(\mathbb{R}), \\ \gamma_{O,\nu}, & \text{if } I \notin \mathcal{R}(\mathbb{R}), \quad O \in \mathcal{R}(\mathbb{R}). \end{cases} \tag{4.63}$$

Then setting $L := \alpha(1+2q)K\gamma$ from relations (4.61) and (4.62), we have that

$$\|f_v\|_{O(\mathbb{R},X)} \leq L \|v\|_{I(\mathbb{R},X)}. \tag{4.64}$$

Now let $v \in C_{0c}(\mathbb{R}, X)$ and $f \in O(\mathbb{R}, X)$ be such that the pair (f, v) satisfies (E_θ). We set $\varphi := f - f_v$, and we have that $\varphi \in O(\mathbb{R}, X)$ and

$$\varphi(t) = \Phi(\sigma(\theta,s), t-s)\varphi(s), \quad \forall t \geq s. \tag{4.65}$$

to Contents, page 1

Let $\varphi_1(t) = P(\sigma(\theta,t))\varphi(t)$, for all $t \in \mathbb{R}$ and let $\varphi_2(t) = (I - P(\sigma(\theta,t)))\varphi(t)$, for all $t \in \mathbb{R}$. Then from (4.65), we obtain that

$$\varphi_k(t) = \Phi(\sigma(\theta,s), t-s)\varphi_k(s), \quad \forall t \geq s, \forall k \in \{1,2\}. \tag{4.66}$$

Let $t_0 \in \mathbb{R}$. From (4.66), it follows that

$$\|\varphi_1(t_0)\| \leq Ke^{-\nu(t_0-s)}\|\varphi_1(s)\| \leq qKe^{-\nu(t_0-s)}\|\varphi(s)\|, \ \forall s \leq t_0. \tag{4.67}$$

Since $\varphi \in O(\mathbb{R},X)$, from Remark 2.12 it follows that $\varphi \in M^1(\mathbb{R},X)$. Then, from (4.67), we have that

$$\begin{aligned} \|\varphi_1(t_0)\| &\leq qK\int_{s-1}^{s} e^{-\nu(t_0-\tau)}\|\varphi(\tau)\|d\tau \leq qKe^{-\nu(t_0-s)}\int_{s-1}^{s}\|\varphi(\tau)\|d\tau \\ &\leq qKe^{-\nu(t_0-s)}\|\varphi\|_{M^1(\mathbb{R},X)}, \quad \forall s \leq t_0. \end{aligned} \tag{4.68}$$

For $s \longrightarrow -\infty$ in (4.68), it follows that $\varphi_1(t_0) = 0$. In addition, from (4.66) we have that

$$\frac{1}{K}e^{\nu(t-t_0)}\|\varphi_2(t_0)\| \leq \|\varphi_2(t)\| \leq (1+q)\|\varphi(t)\|, \quad \forall t \geq t_0. \tag{4.69}$$

This implies that

$$\frac{1}{K}e^{\nu(t-t_0)}\|\varphi_2(t_0)\| \leq (1+q)\int_t^{t+1}\|\varphi(\tau)\|d\tau \leq (1+q)\|\varphi\|_{M^1(\mathbb{R},X)}, \quad \forall t \geq t_0. \tag{4.70}$$

The relation (4.70) shows that

$$\|\varphi_2(t_0)\| \leq K(1+q)e^{-\nu(t-t_0)}\|\varphi\|_{M^1(\mathbb{R},X)}, \quad \forall t \geq t_0. \tag{4.71}$$

For $t \longrightarrow \infty$ in (4.71), it follows that $\varphi_2(t_0) = 0$. This shows that $\varphi(t_0) = \varphi_1(t_0) + \varphi_2(t_0) = 0$. Since $t_0 \in \mathbb{R}$ was arbitrary, we deduce that $\varphi = 0$, so $f = f_v$. Then, from (4.64), we have that

$$\|f\|_{O(\mathbb{R},X)} \leq L\|v\|_{I(\mathbb{R},X)}. \tag{4.72}$$

Taking into account that L does not depend on $\theta \in \Theta$ or on $v \in C_{0c}(\mathbb{R},X)$, we finally conclude that the pair $(O(\mathbb{R},X), I(\mathbb{R},X))$ is uniformly admissible for the system (E_π).

Sufficiency follows from (i). □

Corollary 4.8. *Let $\pi = (\Phi,\sigma)$ be a skew-product flow on $\mathcal{E} = X \times \Theta$ and let V be a Banach function space with $V \in \mathcal{T}(\mathbb{R})$. Then, the following assertions hold:*

(i) *if the pair $(V(\mathbb{R},X), V(\mathbb{R},X))$ is uniformly admissible for the system (E_π), then, π is exponentially dichotomic;*

to Contents, page 1

(ii) *if* $V \in \mathcal{R}(\mathbb{R})$*, then,* π *is exponentially dichotomic if and only if the pair* $(V(\mathbb{R}, X), V(\mathbb{R}, X))$ *is uniformly admissible for the system* (E_π)*.*

Proof. We prove that either $V \in \mathcal{Q}(\mathbb{R})$ or $V \in \mathcal{L}(\mathbb{R})$. Indeed, suppose by contrary that $V \notin \mathcal{Q}(\mathbb{R})$ and $V \notin \mathcal{L}(\mathbb{R})$. Then, $M := \sup_{t>0} F_V(t) < \infty$ and $V \subset L^1(\mathbb{R}, \mathbb{R})$. From $V \subset L^1(\mathbb{R}, \mathbb{R})$, it follows that there is $\gamma > 0$ such that

$$\|v\|_1 \leq \gamma |v|_V, \quad \forall v \in V. \tag{4.73}$$

In particular, from $v = \chi_{[0,t)}$ in relation (4.73), we deduce that

$$t \leq \gamma \left|\chi_{[0,t)}\right|_V = \gamma F_V(t) \leq \gamma M, \quad \forall t > 0, \tag{4.74}$$

which is absurd. This shows that the assumption is false, which shows that either $V \in \mathcal{Q}(\mathbb{R})$ or $V \in \mathcal{L}(\mathbb{R})$. By applying Theorem 4.7, we obtain the conclusion. □

5. Applications and Conclusions

We have seen in the previous section that in the study of the exponential dichotomy of variational equations the classes $\mathcal{Q}(\mathbb{R})$ and, respectively, $\mathcal{L}(\mathbb{R})$ have a crucial role in the identification of the appropriate function spaces in the admissible pair. Moreover, it was also important to point out that it is sufficient to impose conditions either on the input space or on the output space. In this context, the natural question arises if these conditions are indeed necessary and whether our hypotheses may be dropped. The aim of this section is to answer this question. With this purpose, we will present an illustrative example of uniform admissibility, and we will discuss the concrete implications concerning the existence of the exponential dichotomy.

Let X be a Banach space. We denote by $\mathcal{C}_0(\mathbb{R}, X)$ the space of all continuous functions $u : \mathbb{R} \to X$ with $\lim_{t\to\infty} u(t) = \lim_{t\to-\infty} u(t) = 0$, which is a Banach space with respect to the norm

$$|||u||| := \sup_{t\in\mathbb{R}} \|u(t)\|. \tag{5.1}$$

We start with a technical lemma.

Lemma 5.1. *If* O *is a Banach function space with* $O \in \mathcal{T}(\mathbb{R}) \setminus \mathcal{Q}(\mathbb{R})$*, then,* $\mathcal{C}_0(\mathbb{R}, \mathbb{R}) \subset O$*.*

Proof. Let $c := \sup_{t>0} F_O(t)$. Let $u \in \mathcal{C}_0(\mathbb{R}, \mathbb{R})$. Then, there is an unbounded increasing sequence $(t_n) \subset (0, \infty)$ such that $|u(t)| \leq 1/(n+1)$, for all $|t| \geq t_n$ and all $n \in \mathbb{N}$. Setting $u_n = u\chi_{[-t_n,t_n)}$ we have that

$$\left|u_{n+p} - u_n\right|_O \leq \frac{\left|\chi_{[-t_{n+p},-t_n)}\right|_O}{n+1} + \frac{\left|\chi_{[t_n,t_{n+p})}\right|_O}{n+1} \leq \frac{2c}{n+1}, \quad \forall n \in \mathbb{N},\ \forall p \in \mathbb{N}^*. \tag{5.2}$$

From relation (5.2), it follows that the sequence (u_n) is fundamental in O, so this is convergent, that is, there exists $v \in O$ such that $u_n \to v$ in O. According to Remark 2.4(ii),

to Contents, page 1

there exists a subsequence (u_{k_n}) such that $u_{k_n}(t) \to v(t)$ for a.e. $t \in \mathbb{R}$. This implies that $v(t) = u(t)$ for a.e. $t \in \mathbb{R}$, so $v = u$ in O. In conclusion, $u \in O$, and the proof is complete. □

In what follows, we present a concrete situation which illustrates the relevance of the hypotheses on the underlying function spaces considered in the admissible pair, for the study of the dichotomous behavior of skew-product flows.

Example 5.2. Let $X = \mathbb{R} \times \mathbb{R}$ which is a Banach space with respect to the norm $\|(x_1, x_2)\| = |x_1| + |x_2|$. Let $\Theta = \mathbb{R}$ and let $\sigma : \Theta \times \mathbb{R} \to \Theta$, $\sigma(\theta, t) = \theta + t$. We have that σ is a flow on Θ. Let

$$\varphi : \mathbb{R} \longrightarrow (0, \infty), \qquad \varphi(t) = \begin{cases} \dfrac{2}{t+1}, & t \geq 0, \\ 1 + e^{-t}, & t < 0. \end{cases} \tag{5.3}$$

For every $(\theta, t) \in \Theta \times \mathbb{R}_+$, we consider the operator

$$\Phi(\theta, t) : X \longrightarrow X, \quad \Phi(\theta, t)(x_1, x_2) = \left(\frac{\varphi(\theta + t)}{\varphi(\theta)} x_1, e^t x_2 \right). \tag{5.4}$$

It is easy to see that $\pi = (\Phi, \sigma)$ is a skew-product flow.

Now, let O, I be two Banach function spaces with $O, I \in \mathcal{T}(\mathbb{R})$ such that $O \notin \mathcal{Q}(\mathbb{R})$ and $I \notin \mathcal{L}(\mathbb{R})$. It follows that $I \subset L^1(\mathbb{R}, \mathbb{R})$, and, using Lemma 5.1, we obtain that $C_0(\mathbb{R}, \mathbb{R}) \subset O$. Then, there are $\alpha, \beta > 0$ such that

$$\begin{gathered} \|u\|_1 \leq \alpha |u|_I, \quad \forall u \in I, \\ \|u\|_O \leq \beta \|u\|, \quad \forall u \in C_0(\mathbb{R}, \mathbb{R}). \end{gathered} \tag{5.5}$$

Step 1. We prove that the pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ is uniformly admissible for the system (E_π).

Let $\theta \in \Theta$ and let $v = (v_1, v_2) \in C_{0c}(\mathbb{R}, X)$ and let $h > 0$ be such that $\operatorname{supp} v \subset (0, h)$. We consider the function $f : \mathbb{R} \to X$ where $f = (f_1, f_2)$ and

$$f_1(t) = \int_{-\infty}^{t} \frac{\varphi(\theta + t)}{\varphi(\theta + \tau)} v_1(\tau) d\tau, \qquad f_2(t) = -\int_t^{\infty} e^{-(\tau - t)} v_2(\tau) d\tau, \quad \forall t \in \mathbb{R}. \tag{5.6}$$

We have that f is continuous and an easy computation shows that the pair (f, v) satisfies (E_θ). Since $\operatorname{supp} v \subset (0, h)$, we obtain that $f_1(t) = 0$, for all $t \leq 0$ and $f_2(t) = 0$, for all $t \geq h$. From

$$f_1(t) = \varphi(\theta + t) \int_0^h \frac{v_1(\tau)}{\varphi(\theta + \tau)} d\tau, \qquad \forall t \geq h, \tag{5.7}$$

we have that $\lim_{t \to \infty} f_1(t) = 0$. In addition, from

$$f_2(t) = -e^t \int_0^h e^{-\tau} v_2(\tau) d\tau, \quad \forall t \leq 0, \tag{5.8}$$

to Contents, page 1

we deduce that $\lim_{t\to-\infty} f_2(t) = 0$. Thus, we obtain that $f \in \mathcal{C}_0(\mathbb{R}, X)$ so $f \in O(\mathbb{R}, X)$. Moreover, from

$$\begin{aligned} \left|f_1(t)\right| &\le \int_{-\infty}^{t} |v_1(\tau)|d\tau \le \|v_1\|_{L^1(\mathbb{R},\mathbb{R})}, \quad \forall t \in \mathbb{R}, \\ \left|f_2(t)\right| &\le \int_{t}^{\infty} |v_2(\tau)|d\tau \le \|v_2\|_{L^1(\mathbb{R},\mathbb{R})}, \quad \forall t \in \mathbb{R}, \end{aligned} \tag{5.9}$$

it follows that

$$|||f||| \le \|v\|_{L^1(\mathbb{R},X)}. \tag{5.10}$$

From relations (5.5) and (5.10), we obtain that

$$\|f\|_{O(\mathbb{R},X)} \le \alpha\beta\|v\|_{I(\mathbb{R},X)}. \tag{5.11}$$

Let $\tilde{f} \in O(\mathbb{R}, X)$ be such that the pair $(\tilde{f}, v)$ satisfies (E_θ) and let $g = \tilde{f} - f$. Then, $g \in O(\mathbb{R}, X)$ and $g(t) = \Phi(\sigma(\theta, s), t - s)g(s)$, for all $t \ge s$. More exactly, if $g = (g_1, g_2)$, then we have that

$$g_1(t) = \frac{\varphi(\theta + t)}{\varphi(\theta + s)} g_1(s), \quad \forall t \ge s, \tag{5.12}$$

$$g_2(t) = e^{t-s} g_2(s), \quad \forall t \ge s. \tag{5.13}$$

Since $g \in O(\mathbb{R}, X)$ from Remark 2.12, it follows that $g \in M^1(\mathbb{R}, X)$, so $g_1, g_2 \in M^1(\mathbb{R}, \mathbb{R})$.

Let $t_0 \in \mathbb{R}$. For every $s \le t_0$ from relation (5.12), we have that

$$\frac{\left|g_1(t_0)\right|}{\varphi(\theta + t_0)} = \int_{s-1}^{s} \frac{\left|g_1(\tau)\right|}{\varphi(\theta + \tau)} d\tau \le \frac{1}{\varphi(\theta + s)} \int_{s-1}^{s} \left|g_1(\tau)\right| d\tau \le \frac{\|g_1\|_{M^1(\mathbb{R},\mathbb{R})}}{\varphi(\theta + s)}. \tag{5.14}$$

Since $\varphi(r) \to \infty$ as $r \to -\infty$, for $s \to -\infty$ in (5.14), we obtain that $g_1(t_0) = 0$. In addition, for every $t \ge t_0$ from relation (5.13) we have that

$$e^{-t_0}\left|g_2(t_0)\right| = \int_{t}^{t+1} e^{-\tau}\left|g_2(\tau)\right| d\tau \le e^{-t} \int_{t}^{t+1} \left|g_2(\tau)\right| d\tau \le e^{-t}\|g_2\|_{M^1(\mathbb{R},\mathbb{R})}. \tag{5.15}$$

For $t \to \infty$ in (5.15) we deduce that $g_2(t_0) = 0$. So, we obtain that $g(t_0) = 0$. Taking into account that $t_0 \in \mathbb{R}$ was arbitrary it follows that $g = 0$. This implies that $\tilde{f} = f$. Then, from relation (5.11) we have that

$$\left\|\tilde{f}\right\|_{O(\mathbb{R},X)} \le \alpha\beta\|v\|_{I(\mathbb{R},X)}. \tag{5.16}$$

We set $L = \alpha\beta$, and, taking into account that L does not depend on θ or v, we conclude that the pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ is uniformly admissible for the system (E_π).

Step 2. We prove that π is not exponentially dichotomic. Suppose by contrary that π is exponentially dichotomic with respect to the family of projections $\{P(\theta)\}_{\theta\in\Theta}$ and let $K, \nu > 0$ be two constants given by Definition 4.1. In this case, according to Proposition 2.1 from [18] we have that

$$\operatorname{Im} P(\theta) = \{x \in X : \Phi(\theta, t)x \longrightarrow 0 \text{ as } t \longrightarrow \infty\}, \quad \forall\theta \in \Theta. \tag{5.17}$$

This characterization implies that $\operatorname{Im} P(\theta) = \mathbb{R} \times \{0\}$, for all $\theta \in \Theta$. Then, from

$$\|\Phi(\theta, t)x\| \le Ke^{-\nu t}\|x\|, \quad \forall t \ge 0,\ \forall x \in \operatorname{Im} P(\theta),\ \forall\theta \in \Theta, \tag{5.18}$$

we obtain that

$$\frac{\varphi(\theta + t)}{\varphi(\theta)}|x_1| \le Ke^{-\nu t}|x_1|, \quad \forall x_1 \in \mathbb{R},\ \forall t \ge 0,\ \forall\theta \in \Theta, \tag{5.19}$$

which shows that

$$\frac{\varphi(\theta + t)}{\varphi(\theta)} \le Ke^{-\nu t}, \quad \forall t \ge 0,\ \forall\theta \in \Theta. \tag{5.20}$$

In particular, for $\theta = 0$, from (5.20), we have that

$$\frac{1}{t + 1} \le Ke^{-\nu t}, \quad \forall t \ge 0, \tag{5.21}$$

which is absurd. This shows that the assumption is false, so π is not exponentially dichotomic.

Remark 5.3. The above example shows that if I, O are two Banach function spaces from the class $\mathcal{T}(\mathbb{R})$ such that $O \notin \mathcal{Q}(\mathbb{R})$ and $I \notin \mathcal{L}(\mathbb{R})$, then the uniform admissibility of the pair $(O(\mathbb{R}, X), I(\mathbb{R}, X))$ for the system (E_π) *does not imply* the existence of the exponential dichotomy of π. This shows that the hypotheses of the main result from the previous section are indeed necessary and emphasizes the fact that in the study of the exponential dichotomy in terms of the uniform admissibility at least one of the output space or the input space should belong to, respectively, $\mathcal{Q}(\mathbb{R})$ or $\mathcal{L}(\mathbb{R})$.

Finally, we complete our study with several consequences of the main result, which will point out some interesting conclusions for some usual classes of spaces often used in control-type problems arising in qualitative theory of dynamical systems. We will also show that, in our approach, the input space can be successively minimized, and we will discuss several optimization directions concerning the admissibility-type techniques.

Remark 5.4. The input-output characterizations for the asymptotic properties of systems have a wider applicability area if the input space is as small as possible and the output space is

to Contents, page 1

very general. In our main result, given by Theorem 4.7, the input functions belong to the space $C_{0c}(\mathbb{R}, X)$ while the output space is a general Banach function space. By analyzing condition (ii) from Definition 3.6, we observe that the input-output characterization given by Theorem 4.7 becomes more flexible and provides a more competitive applicability spectrum when the norm on the input space is larger.

Another interesting aspect that must be noted is that the class $\mathcal{T}(\mathbb{R})$ is closed to finite intersections. Indeed, if $I_1, \ldots, I_n \in \mathcal{T}(\mathbb{R})$, then we may define $I := I_1 \cap I_2 \cap \cdots I_n$ with respect to the norm

$$|u|_I := \max\{|u|_{I_1}, |u|_{I_2}, \ldots, |u|_{I_n}\}, \tag{5.22}$$

which is a Banach function space which belongs to $\mathcal{T}(\mathbb{R})$. So, taking as input space a Banach function space which is obtained as an intersection of Banach function spaces from the class $\mathcal{T}(\mathbb{R})$ we will have a "larger" norm in our admissibility condition, and, thus the estimation will be more permissive and more general.

As a consequence of the aspects presented in the above remark we deduce the following corollaries.

Corollary 5.5. *Let $\pi = (\Phi, \sigma)$ be a skew-product flow on $X \times \Theta$. Let O_φ be an Orlicz space with $0 < \varphi(t) < \infty$, for all $t > 0$. Let $n \in \mathbb{N}^*$, let $O_{\varphi_1}, \ldots, O_{\varphi_n}$ be Orlicz spaces such that $\varphi_k(1) < \infty$, for all $k \in \{1, \ldots, n\}$ and let $I := O_{\varphi_1}(\mathbb{R}, \mathbb{R}) \cap \cdots \cap O_{\varphi_n}(\mathbb{R}, \mathbb{R}) \cap O_\varphi(\mathbb{R}, \mathbb{R})$. Then, π is exponentially dichotomic if and only if the pair $(O_\varphi(\mathbb{R}, X), I(\mathbb{R}, X))$ is uniformly admissible for the system (E_π).*

Proof. From Lemma 2.15 and Remark 2.20, it follows that $O_\varphi \in \mathcal{Q}(\mathbb{R}) \cap \mathcal{R}(\mathbb{R})$. By applying Theorem 4.7, the proof is complete. □

Corollary 5.6. *Let $\pi = (\Phi, \sigma)$ be a skew-product flow on $X \times \Theta$ and let $p \in [1, \infty)$. Let $n \in \mathbb{N}^*, q_1, \ldots, q_n \in [1, \infty]$ and $I = L^{q_1}(\mathbb{R}, \mathbb{R}) \cap \cdots \cap L^{q_n}(\mathbb{R}, \mathbb{R}) \cap L^p(\mathbb{R}, \mathbb{R})$. Then, π is exponentially dichotomic if and only if the pair $(L^p(\mathbb{R}, X), I(\mathbb{R}, X))$ is admissible for the system (E_π).*

Proof. This follows from Corollary 5.5. □

Corollary 5.7. *Let $\pi = (\Phi, \sigma)$ be a skew-product flow on $X \times \Theta$ and let $p \in (1, \infty]$. Let $n \in \mathbb{N}^*, q_1, \ldots, q_n \in (1, \infty]$ and $I = L^{q_1}(\mathbb{R}, \mathbb{R}) \cap \cdots \cap L^{q_n}(\mathbb{R}, \mathbb{R}) \cap L^p(\mathbb{R}, \mathbb{R})$. Then π is exponentially dichotomic if and only if the pair $(L^p(\mathbb{R}, X), I(\mathbb{R}, X))$ is uniformly admissible for the system (E_π).*

Proof. This follows from Theorem 4.7 by observing that $I \in \mathcal{L}(\mathbb{R})$. □

Remark 5.8. According to Remark 2.12, the largest space from the class $\mathcal{T}(\mathbb{R})$ is $M^1(\mathbb{R}, \mathbb{R})$. Thus, considering the output space $M^1(\mathbb{R}, \mathbb{R})$, in order to obtain optimal input-output characterizations for exponential dichotomy in terms of admissibility, it is sufficient to work with smaller and smaller input spaces.

Corollary 5.9. *Let $\pi = (\Phi, \sigma)$ be a skew-product flow on $X \times \Theta$. Let $n \in \mathbb{N}^*$, $q_1, \ldots, q_n \in (1, \infty]$ and $I = L^{q_1}(\mathbb{R}, \mathbb{R}) \cap \cdots \cap L^{q_n}(\mathbb{R}, \mathbb{R})$. Then, π is exponentially dichotomic if and only if the pair $(M^1(\mathbb{R}, X), I(\mathbb{R}, X))$ is uniformly admissible for the system (E_π).*

to Contents, page 1

Proof. We observe that $I \in \mathcal{L}(\mathbb{R})$, and, from Remark 2.12, we have that $I \subset M^1(\mathbb{R}, \mathbb{R})$. By applying Theorem 4.7, we obtain the conclusion. □

Acknowledgment

This work was supported by CNCSIS-UEFISCDI, project number PN II-IDEI code 1080/2008 no. 508/2009.

References

[1] E. Braverman and S. Zhukovskiy, "The problem of a lazy tester, or exponential dichotomy for impulsive differential equations revisited," *Nonlinear Analysis: Hybrid Systems*, vol. 2, no. 3, pp. 971–979, 2008.

[2] C. Cuevas and C. Vidal, "Discrete dichotomies and asymptotic behavior for abstract retarded functional difference equations in phase space," *Journal of Difference Equations and Applications*, vol. 8, no. 7, pp. 603–640, 2002.

[3] S.-N. Chow and H. Leiva, "Existence and roughness of the exponential dichotomy for skew-product semiflow in Banach spaces," *Journal of Differential Equations*, vol. 120, no. 2, pp. 429–477, 1995.

[4] S.-N. Chow and H. Leiva, "Two definitions of exponential dichotomy for skew-product semiflow in Banach spaces," *Proceedings of the American Mathematical Society*, vol. 124, no. 4, pp. 1071–1081, 1996.

[5] S.-N. Chow and H. Leiva, "Unbounded perturbation of the exponential dichotomy for evolution equations," *Journal of Differential Equations*, vol. 129, no. 2, pp. 509–531, 1996.

[6] J. H. Liu, G. M. N'Guérékata, and N. Van Minh, *Topics on Stability and Periodicity in Abstract Differential Equations*, vol. 6 of *Series on Concrete and Applicable Mathematics*, World Scientific, Hackensack, NJ, USA, 2008.

[7] N. Van Minh, F. Räbiger, and R. Schnaubelt, "Exponential stability, exponential expansiveness, and exponential dichotomy of evolution equations on the half-line," *Integral Equations and Operator Theory*, vol. 32, no. 3, pp. 332–353, 1998.

[8] N. Van Minh, G. M. N'Guérékata, and R. Yuan, *Lectures on the Asymptotic Behavior of Solutions of Differential Equations*, Nova Science, New York, NY, USA, 2008.

[9] K. J. Palmer, "Exponential dichotomies and Fredholm operators," *Proceedings of the American Mathematical Society*, vol. 104, no. 1, pp. 149–156, 1988.

[10] K. J. Palmer, *Shadowing in Dynamical Systems*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.

[11] O. Perron, "Die stabilitätsfrage bei differentialgleichungen," *Mathematische Zeitschrift*, vol. 32, no. 1, pp. 703–728, 1930.

[12] M. Pituk, "A criterion for the exponential stability of linear difference equations," *Applied Mathematics Letters*, vol. 17, no. 7, pp. 779–783, 2004.

[13] H. M. Rodrigues and J. G. Ruas-Filho, "Evolution equations: dichotomies and the Fredholm alternative for bounded solutions," *Journal of Differential Equations*, vol. 119, no. 2, pp. 263–283, 1995.

[14] B. Sasu and A. L. Sasu, "Input-output conditions for the asymptotic behavior of linear skew-product flows and applications," *Communications on Pure and Applied Analysis*, vol. 5, no. 3, pp. 551–569, 2006.

[15] B. Sasu, "New criteria for exponential expansiveness of variational difference equations," *Journal of Mathematical Analysis and Applications*, vol. 327, no. 1, pp. 287–297, 2007.

[16] B. Sasu, "Robust stability and stability radius for variational control systems," *Abstract and Applied Analysis*, vol. 2008, Article ID 381791, 29 pages, 2008.

[17] B. Sasu, "On dichotomous behavior of variational difference equations and applications," *Discrete Dynamics in Nature and Society*, vol. 2009, Article ID 140369, 16 pages, 2009.

[18] B. Sasu, "Integral conditions for exponential dichotomy: a nonlinear approach," *Bulletin des Sciences Mathématiques*, vol. 134, no. 3, pp. 235–246, 2010.

[19] C. Foias, G. R. Sell, and R. Temam, "Inertial manifolds for nonlinear evolutionary equations," *Journal of Differential Equations*, vol. 73, no. 2, pp. 309–353, 1988.

to Contents, page 1

[20] W. A. Coppel, *Dichotomies in Stability Theory*, vol. 629 of *Lecture Notes in Mathematics*, Springer, Berlin, Germany, 1978.

[21] J. L. Daleckii and M. G. Krein, *Stability of Solutions of Differential Equations in Banach Space*, American Mathematical Society, Providence, RI, USA, 1974.

[22] J. L. Massera and J. J. Schäffer, *Linear Differential Equations and Function Spaces*, vol. 21 of *Pure and Applied Mathematics*, Academic Press, New York, NY, USA, 1966.

[23] L. Berezansky and E. Braverman, "On exponential dichotomy, Bohl-Perron type theorems and stability of difference equations," *Journal of Mathematical Analysis and Applications*, vol. 304, no. 2, pp. 511–530, 2005.

[24] A. D. Maĭzel', "On stability of solutions of systems of differential equations," *Trudy Ural'skogo Politekhnicheskogo Instituta*, vol. 51, pp. 20–50, 1954 (Russian).

[25] P. Meyer-Nieberg, *Banach Lattices*, Universitext, Springer, Berlin, Germany, 1991.

[26] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, NY, USA, 3rd edition, 1987.

[27] C. Bennett and R. Sharpley, *Interpolation of Operators*, vol. 129 of *Pure and Applied Mathematics*, Academic Press, Boston, Mass, USA, 1988.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 182827, 16 pages
doi:10.1155/2011/182827

Research Article

Two-Parametric Conditionally Oscillatory Half-Linear Differential Equations

Ondřej Došlý[1] and Simona Fišnarová[2]

[1] *Department of Mathematics and Statistics, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic*
[2] *Department of Mathematics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic*

Correspondence should be addressed to Ondřej Došlý, dosly@math.muni.cz

Received 2 November 2010; Accepted 5 January 2011

Academic Editor: Miroslava Růžičková

We study perturbations of the nonoscillatory half-linear differential equation $(r(t)\Phi(x'))' + c(t)\Phi(x) = 0$, $\Phi(x) := |x|^{p-2}x$, $p > 1$. We find explicit formulas for the functions $\hat{r}$, $\hat{c}$ such that the equation $[(r(t) + \lambda\hat{r}(t))\Phi(x')]' + [c(t) + \mu\hat{c}(t)]\Phi(x) = 0$ is conditionally oscillatory, that is, there exists a constant γ such that the previous equation is oscillatory if $\mu - \lambda > \gamma$ and nonoscillatory if $\mu - \lambda < \gamma$. The obtained results extend the previous results concerning two-parametric perturbations of the half-linear Euler differential equation.

1. Introduction

Conditionally oscillatory equations play an important role in the oscillation theory of the Sturm-Liouville second-order differential equation

$$\left(r(t)x'\right)' + c(t)x = 0, \tag{1.1}$$

with positive continuous functions r, c. Equation (1.1) with λc instead of c is said to be *conditionally oscillatory* if there exists $\lambda_0 > 0$, the so-called *oscillation constant* of (1.1), such that this equation is oscillatory for $\lambda > \lambda_0$ and nonoscillatory for $\lambda < \lambda_0$. A typical example of a conditionally oscillatory equation is the Euler differential equation

$$x'' + \frac{\lambda}{t^2}x = 0, \tag{1.2}$$

to Contents, page 1

which has the oscillation constant $\lambda_0 = 1/4$ as can be verified by a direct computation when looking for solutions of (1.2) in the form $x(t) = t^{\alpha}$. This leads to the classical Kneser (non)oscillation criterion which states that (1.1) with $r(t) \equiv 1$ is oscillatory provided

$$\liminf_{t\to\infty} t^2 c(t) > \frac{1}{4}, \tag{1.3}$$

and nonoscillatory if

$$\limsup_{t\to\infty} t^2 c(t) < \frac{1}{4}. \tag{1.4}$$

This shows that the potential $c(t) = t^{-2}$ is the border line between oscillation and nonoscillation. Note that the concept of conditional oscillation of (1.1) was introduced in [1].

The linear oscillation theory extends almost verbatim to the half-linear differential equation

$$\left(r(t)\Phi\left(x'\right)\right)' + c(t)\Phi(x) = 0, \quad \Phi(x) := |x|^{p-2}x, \ p > 1, \tag{1.5}$$

including the definition of conditional oscillation. The half-linear version of Euler equation (1.2) is the equation

$$\left(\Phi\left(x'\right)\right)' + \frac{\lambda}{t^p}\Phi(x) = 0, \tag{1.6}$$

which has the oscillation constant $\lambda_0 = \gamma_p := ((p-1)/p)^p$, and (non)oscillation criteria (1.3), (1.4) extend in a natural way to (1.5) with $r(t) \equiv 1$. A complementary concept to the conditional oscillation is the concept of strong (non)oscillation. Equation (1.5) with λc instead of c is said to be *strongly (non)oscillatory* if it is (non)oscillatory for every $\lambda > 0$. Sometimes, strongly oscillatory equations are regarded as conditionally oscillatory with the oscillation constant $\lambda_0 = 0$ and strongly nonoscillatory as conditionally oscillatory with the oscillation constant $\lambda_0 = \infty$. We refer to [2] for results along this line.

In our paper, we are motivated by a statement presented in [3, 4], where the two-parametric perturbation of the Euler differential equation with the critical coefficient

$$\left(\Phi\left(x'\right)\right)' + \frac{\gamma_p}{t^p}\Phi(x) = 0 \tag{1.7}$$

is investigated. It is shown there that the equation

$$\left[\left(1 + \frac{\lambda}{\log^2 t}\right)\Phi\left(x'\right)\right]' + \left[\frac{\gamma_p}{t^p} + \frac{\mu}{t^p \log^2 t}\right]\Phi(x) = 0 \tag{1.8}$$

to Contents, page 1

is oscillatory if $\mu - \gamma_p\lambda > \mu_p := (1/2)((p-1)/p)^{p-1}$ and nonoscillatory in the opposite case. Note that an important role in proving the results of [4] is played by the fact that we know explicitly the solution $h(t) = t^{(p-1)/p}$ of (1.7).

Here, we treat the problem of conditional oscillation in the following general setting. We suppose that (1.5) is nonoscillatory and that h is its eventually positive solution. We find explicit formulas for the functions $\widehat{r}$, $\widehat{c}$ such that the equation

$$\left[(r(t) + \lambda\widehat{r}(t))\Phi(x')\right]' + \left[c(t) + \mu\widehat{c}(t)\right]\Phi(x) = 0 \tag{1.9}$$

is conditionally oscillatory, that is, there exists a constant γ such that (1.9) is oscillatory if $\mu - \lambda > \gamma$ and nonoscillatory if $\mu - \lambda < \gamma$.

The setup of the paper is as follows. In the next section, we present some statements of the half-linear oscillation theory. Section 3 is devoted to the so-called modified Riccati equation associated with (1.5) and (1.9). The main result of the paper, the construction of the functions $\widehat{r}$, $\widehat{c}$ such that (1.9) is two-parametric conditionally oscillatory, is presented in Section 4.

2. Auxiliary Results

As we have already mentioned in the previous section, the linear oscillation theory extends almost verbatim to half-linear equation (1.5). The word "almost" reflects the fact that not all linear methods can be extended to (1.5), some results for (1.5) are the same as those for (1.1), but to prove them, one has to use different methods than in the linear case. A typical method of this kind is the following transformation formula. If $f(t) \neq 0$ is a sufficiently smooth function and functions x, y are related by the formula $x = f(t)y$, then we have the identity

$$f(t)\left[(r(t)x')' + c(t)x\right] = (R(t)y')' + C(t)y, \tag{2.1}$$

where

$$R(t) = r(t)f^2(t), \qquad C(t) = f(t)\left[(r(t)f'(t))' + c(t)f(t)\right]. \tag{2.2}$$

In particular, x is a solution of (1.1) if and only if y is a solution of the equation $(Ry')' + Cy = 0$. The transformation identity (2.1) *does not extend* to (1.5).

To illustrate the meaning of this fact in the conditional oscillation of (1.1) and (1.5), suppose that (1.1) is nonoscillatory and let h be its so-called *principal solution* (see [5, Chapter XI]), that is, a solution such that $\int^\infty r^{-1}(t)h^{-2}(t)\mathrm{d}t = \infty$. We would like to find a function $\widehat{c}$ such that the equation

$$(r(t)x')' + (c(t) + \mu\widehat{c}(t))x = 0 \tag{2.3}$$

is conditionally oscillatory and to find its oscillation constant. The transformation $x = h(t)y$ transforms (1.1) into the one term equation $(r(t)h^2(t)y')' = 0$ and the transformation of independent variable $s = \int^t r^{-1}(\tau)h^{-2}(\tau)\mathrm{d}\tau$ further to the equation $\mathrm{d}^2y/\mathrm{d}s^2 = 0$. Now,

to Contents, page 1

from (1.2), we know that the "right" perturbation term in the last equation is $1/s^2$ with the oscillation constant $1/4$. Substituting back for s, we get the conditionally oscillatory equation

$$(R(t)y')' + \frac{\mu}{R(t)\left(\int^t R^{-1}(s)\mathrm{d}s\right)^2} y = 0, \quad R(t) = r(t)h^2(t), \tag{2.4}$$

and the back transformation $y = h^{-1}(t)x$ results in the conditionally oscillatory equation

$$(r(t)x')' + \left[c(t) + \frac{\mu}{h^2(t)R(t)\left(\int^t R^{-1}(s)\mathrm{d}s\right)^2}\right] x = 0, \tag{2.5}$$

with the oscillation constant $\mu_0 = 1/4$. The previous result is one of the main statements of [6], but it was proved there by a different method.

In the next section, we will show how to modify this method to be applicable to half-linear equations. At this moment, we present the result of [7] with the classical (i.e., one parametric) conditional oscillation of (1.5). Let h be a positive solution of (1.5) such that $h'(t) \neq 0$ for large t. We denote

$$R(t) := r(t)h^2(t)|h'(t)|^{p-2}, \qquad G(t) := r(t)h(t)\Phi(h'(t)), \tag{2.6}$$

$$\widehat{c}(t) = \frac{1}{|h(t)|^p R(t)\left(\int^t R^{-1}(s)\mathrm{d}s\right)^2}. \tag{2.7}$$

Theorem 2.1. *Suppose that* (1.5) *possesses a nonoscillatory solution h such that $h'(t) \neq 0$ for large t, and R, G are given by* (2.6)*. If*

$$\int^\infty \frac{\mathrm{d}t}{R(t)} = \infty, \qquad \liminf_{t\to\infty} |G(t)| > 0, \tag{2.8}$$

then the equation

$$(r(t)\Phi(x'))' + [c(t) + \mu\widehat{c}(t)]\Phi(x) = 0 \tag{2.9}$$

is conditionally oscillatory, and its oscillation constant is $\mu_0 = 1/2q$, where q is the conjugate exponent to p, that is, $1/p + 1/q = 1$.

Note that in the linear case $p = 2$, the function $f(t) = h(t)\sqrt{\int^t r^{-1}(\tau)h^{-2}(\tau)\mathrm{d}\tau}$ is a solution of (2.9) with $\mu = \mu_0 = 1/4$. In the general half-linear case, an explicit solution of (2.9) is no longer known, but we are able to "estimate" this solution. The next statement, which is also taken from [7], presents a result along this line.

to Contents, page 1

Theorem 2.2. *Suppose that* (2.8) *holds and let* $f(t) = h(t)(\int^t R^{-1}(s)ds)^{1/p}$, *then a solution of* (2.9) *with* $\mu = 1/2q$ *is of the form*

$$x(t) = f(t)\left(1 + O\left(\left(\int^t R^{-1}(s)ds\right)^{-1}\right)\right), \tag{2.10}$$

and (suppressing the argument t)

$$\begin{aligned} &f\left[(r\Phi(f'))' + \left(c + \frac{1}{2qh^p R\left(\int^t R^{-1}\right)^2}\right)\Phi(f)\right] \\ &\quad = -\frac{(p-1)(p-2)G'}{G^2\left(\int^t R^{-1}\right)} - \frac{(p-1)(p-2)}{3p^3G^3\left(\int^t R^{-1}\right)^2}\left[(p-3)G' + 2pr\left|h'\right|^p\right] \\ &\qquad + O\left(G^{-3}\left(\int^t R^{-1}\right)^{-3}\right)\left[\frac{G'}{pG^2} - \frac{(p^3-4p^2+11p-6)h'}{2p^3h} - \frac{1}{qR\left(\int^t R^{-1}\right)}\right], \end{aligned} \tag{2.11}$$

as $t \to \infty$.

The last statement presented in this section is the so-called *reciprocity principle*. Let x be a solution of (1.5) and let $u := r\Phi(x')$ be its *quasiderivative*, then u is a solution of the reciprocal equation

$$\left(c^{1-q}(t)\Phi^{-1}(u')\right)' + r^{1-q}(t)\Phi^{-1}(u) = 0, \tag{2.12}$$

where $\Phi^{-1}(u) = |u|^{q-2}u$ is the inverse function of Φ.

3. Modified Riccati Equation

Suppose that λ and $\hat{r}$ in (1.9) are such that $r(t) + \lambda\hat{r}(t) > 0$. Let $x(t) \neq 0$ in an interval I be a solution of (1.9), and let $w = (r + \lambda\hat{r})\Phi(x'/x)$. Then, w solves in I the "standard" Riccati equation

$$w' + c(t) + \mu\hat{c}(t) + (p-1)\left[r(t) + \lambda\hat{r}(t)\right]^{1-q}|w|^q = 0. \tag{3.1}$$

More precisely, the following statement holds.

Lemma 3.1 ([8, Theorem 2.2.1]). *The following statements are equivalent:*

(i) *equation* (1.9) *is nonoscillatory;*

(ii) *equation* (3.1) *has a solution on an interval* $[T, \infty)$*;*

to Contents, page 1

(iii) *there exists a continuously differentiable function* w *such that*

$$w' + c(t) + \mu\hat{c}(t) + (p-1)[r(t) + \lambda\hat{r}(t)]^{1-q}(t)|w|^q \leq 0 \tag{3.2}$$

on an interval $[T, \infty)$.

In the linear case, if x is a solution of (1.1), $x = f(t)y$, and $v = rf^2y'/y$ is the Riccati variable corresponding to the equation on the right-hand side in (2.1), then $v = f^2(w - w_f)$ where $w = rx'/x$, $w_f = rf'/f$. This suggests to investigate the function $v = f^p(w - w_f)$ in the half-linear case, and this leads to the *modified Riccati equation* introduced in the next statement which is taken from [4] with a modification from [3].

Lemma 3.2. *Suppose that* f *is a positive differentiable function,* $w_f = (r + \lambda\hat{r})\Phi(f'/f)$, *and* w *is a continuously differentiable function, and put* $v = f^p(w - w_f)$, *then the following identity holds:*

$$\begin{aligned} &f^p(t)\Big[w' + c(t) + \mu\hat{c}(t) + (p-1)(r(t) + \lambda\hat{r}(t))^{1-q}|w|^q\Big] \\ &\quad = v' + f(t)\Big[\ell(f(t)) + \hat{\ell}(f(t))\Big] + (p-1)(r(t) + \lambda\hat{r}(t))^{1-q}f^{-q}(t)\mathcal{G}(t,v), \end{aligned} \tag{3.3}$$

where

$$\ell(f) = (r(t)\Phi(f'))' + c(t)\Phi(f), \qquad \hat{\ell}(f) = \lambda(\hat{r}(t)\Phi(f'))' + \mu\hat{c}(t)\Phi(f), \tag{3.4}$$

$$\mathcal{G}(t,v) = |v + \Omega(t)|^q - q\Phi^{-1}(\Omega(t))v - |\Omega(t)|^q, \quad \Omega := (r + \lambda\hat{r})f\Phi(f'). \tag{3.5}$$

In particular, if w *is a solution of* (3.1), *then* v *is a solution of the modified Riccati equation*

$$v' + f(t)\Big[\ell(f(t)) + \hat{\ell}(f(t))\Big] + (p-1)(r(t) + \lambda\hat{r}(t))^{1-q}f^{-q}(t)\mathcal{G}(t,v) = 0. \tag{3.6}$$

Conversely, if v *is a solution of* (3.6), *then* $w = w_f + f^{-p}v$ *is a solution of* (3.1).

Observe that in case $f \equiv 1$, the modified Riccati equation (3.6) reduces to the standard Riccati equation (3.1).

Next, we will investigate the function $\mathcal{G}$ in (3.5). First, we present a result from [4, Lemmas 5 and 6].

Lemma 3.3. *The function* $\mathcal{G}$ *defined in* (3.5) *has the following properties.*

(i) $\mathcal{G}(t,v) \geq 0$ *with the equality if and only if* $v = 0$.

(ii) *If* $q \geq 2$, *one has the inequality*

$$\mathcal{G}(t,v) \geq \frac{q}{2}|\Omega(t)|^{q-2}v^2. \tag{3.7}$$

Now, we concentrate on an estimate of the function $\mathcal{G}$ in case $q < 2$.

to Contents, page 1

Lemma 3.4. *Suppose that* $q < 2$ *and* $\lim_{t\to\infty}|\Omega(t)| = \infty$, *then there is a constant* $\beta > 0$ *such that for* $v \in (-\infty, -v_0]$, $v_0 > 0$, *and large* t

$$\mathcal{G}(t,v) \geq \beta|\Omega(t)|^{q-2}|v|^q. \tag{3.8}$$

Proof. Consider the function

$$\mathcal{H}(t,v) = \begin{cases} \dfrac{\mathcal{G}(t,v)}{|v|^q}, & \text{for } v \neq 0, \\ 0, & \text{for } v = 0. \end{cases} \tag{3.9}$$

First of all,

$$\lim_{v\to\pm\infty}\mathcal{H}(t,v) = 1, \qquad \lim_{v\to 0}\mathcal{H}(t,v) = 0. \tag{3.10}$$

Now, we compute local extrema of $\mathcal{H}$ with respect to v. We have (suppressing the argument t)

$$\begin{aligned} \mathcal{H}_v &= \frac{1}{|v|^{2q}}\left\{\left[q\Phi^{-1}(v+\Omega) - q\Phi^{-1}(\Omega)\right]|v|^q - q\Phi^{-1}(v)\left[|v+\Omega|^q - q\Phi^{-1}(\Omega)v - |\Omega|^q\right]\right\} \\ &= \frac{q}{v^2\Phi^{-1}(v)}\left\{v\Phi^{-1}(v+\Omega) - v\Phi^{-1}(\Omega) - |v+\Omega|^q + q\Phi^{-1}(\Omega)v + |\Omega|^q\right\} \\ &= \frac{q}{v^2\Phi^{-1}(v)}\left\{-\Omega\Phi^{-1}(v+\Omega) + (q-1)\Phi^{-1}(\Omega)v + |\Omega|^q\right\}. \end{aligned} \tag{3.11}$$

Denote $\mathcal{N}(v)$ the function in braces on the last line of the previous computation. We have $\mathcal{N}(0) = 0$,

$$\begin{aligned} \mathcal{N}'(v) &= -(q-1)\Omega|v+\Omega|^{q-2} + (q-1)\Phi^{-1}(\Omega) \\ &= (q-1)\Omega\left[-|v+\Omega|^{q-2} + |\Omega|^{q-2}\right] \\ &= 0 \end{aligned} \tag{3.12}$$

if and only if $v = 0$ and $v = -2\Omega$, and

$$\mathcal{N}''(v) = -(q-1)(q-2)\Omega|v+\Omega|^{q-3}\,\mathrm{sgn}(v+\Omega). \tag{3.13}$$

This means that $v = 0$ is the local minimum and $v = -2\Omega$ is the local maximum of the function $\mathcal{N}$. Using this result, an examination of the graph of the function $\mathcal{H}$ shows that this function has the local minimum at $v = 0$ and a local maximum in the interval $(-\infty, -2\Omega)$ if $\Omega > 0$,

to Contents, page 1

and this maximum is in $(-2\Omega, \infty)$ if $\Omega < 0$. Next, denote v^* the value for which $\mathcal{H}(t, v^*) = 1$. Consequently, for any $v_0 > 0$, it follows from (3.10) that

$$\inf_{v\in(-\infty,-v_0]} \mathcal{H}(t,v) = \mathcal{H}(t,-\overline{v}) = \frac{1}{|\overline{v}|^q}\left[|\Omega-\overline{v}|^q + q\overline{v}\Phi^{-1}(\Omega) - |\Omega|^q\right], \tag{3.14}$$

where

$$\overline{v} = \begin{cases} -v^* & \text{if } -v_0 < v^* < 0, \\ v_0 & \text{otherwise.} \end{cases} \tag{3.15}$$

Next, we want to investigate the dependence of this infimum on Ω when $|\Omega| \to \infty$. To this end, we investigate the function $F(x) = |x-a|^q + qa\Phi^{-1}(x) - |x|^q$ for $x \to \pm\infty$, $a \in \mathbb{R}$ being a parameter. We have (using the expansion formula for $(1+x)^\alpha$)

$$\begin{aligned} F(x) &= \Phi^{-1}(x)\left\{\frac{|x-a|^q - |x|^q}{\Phi^{-1}(x)} + qa\right\} = \Phi^{-1}(x)\left\{x\left[\left(1-\frac{a}{x}\right)^q - 1\right] + qa\right\} \\ &= \Phi^{-1}(x)\left(\binom{q}{2}\frac{a^2}{x} + o\left(x^{-1}\right)\right) = a^2\binom{q}{2}|x|^{q-2}(1+o(1)), \end{aligned} \tag{3.16}$$

as $|x| \to \infty$. Consequently, if $\lim_{t\to\infty}|\Omega(t)| = \infty$, there exists a constant $\beta > 0$ such that (3.8) holds. □

Now, we are ready to formulate a complement of [9, Theorem 2] which is presented in that paper under the assumption that the function Ω is bounded.

Theorem 3.5. *Let f be a positive continuously differentiable function such that $f'(t) \neq 0$ for large t. Suppose that $\int^\infty \mathcal{R}^{-1}(t)\mathrm{d}t = \infty$, where $\mathcal{R} = (r+\lambda\hat{r})f^2|f'|^{p-2}$, $C(t) \geq 0$ for large t, and $\lim_{t\to\infty}|\Omega(t)| = \infty$, then all possible proper solutions (i.e., solutions which exist on some interval of the form $[T,\infty)$) of the equation*

$$v' + C(t) + (p-1)(r(t)+\lambda\hat{r}(t))^{1-q}f^{-q}(t)\mathcal{G}(t,v) = 0 \tag{3.17}$$

are nonnegative.

Proof. First consider the case $q < 2$. Let $v_0 > 0$ be arbitrary. By Lemma 3.4, there exists $T_0 \in \mathbb{R}$ and $\beta > 0$ such that for $t \geq T_0$ and $v \in (-\infty, -v_0]$,

$$(p-1)(r+\lambda\hat{r})^{1-q}f^{-q}\mathcal{G}(t,v) \geq \beta(p-1)(r+\lambda\hat{r})^{1-q}f^{-q}|\Omega|^{q-2}|v|^q = (p-1)\beta\frac{|v|^q}{\mathcal{R}}. \tag{3.18}$$

Suppose that v is the solution of (3.17) such that $v(t_0) = -v_0$ for some $t_0 \geq T_0$, then

$$v' + C(t) + (p-1)\beta\frac{|v|^q}{\mathcal{R}(t)} \leq 0, \tag{3.19}$$

to Contents, page 1

for $t \ge t_0$ for which the solution v exists. Now, we use the same argument as in the proof of Theorem 2 in [9]. Consider the equation

$$z' + C(t) + (p-1)\beta\frac{|z|^q}{\mathcal{R}(t)} = 0. \tag{3.20}$$

This is the standard Riccati equation corresponding to the half-linear equation

$$\left(\mathcal{R}^{p-1}(t)\Phi(x')\right)' + \beta^{p-1}C(t)\Phi(x) = 0. \tag{3.21}$$

Assumptions of theorem imply, by [8, Corollary 4.2.1], that all proper solutions of (3.20) are nonnegative. It means that any solution of (3.20) which starts with a negative initial condition blows down to $-\infty$ in a finite time. Inequality (3.19) implies that if z is the solution of (3.20) satisfying $z(t_0) = v(t_0) = -v_0$, that is, z starts with the same initial value as the solution v of (3.17), then v decreases faster than z. In particular, if z blows down to $-\infty$ at a finite time, then v does as well. This means that all proper solutions of (3.17), if any, are nonnegative.

In case $q \ge 2$, we proceed in a similar way. We use (3.7) and we compare (3.17) with the equation

$$z' + C(t) + \frac{p}{2}\frac{z^2}{\mathcal{R}(t)} = 0, \tag{3.22}$$

which is the standard Riccati equation corresponding to the linear equation

$$(\mathcal{R}(t)x')' + \frac{p}{2}C(t)x = 0. \tag{3.23}$$

Then, reasoning in the same way as in case $q < 2$, we obtain the conclusion that all proper solutions of (3.17) are nonnegative also in this case. □

4. Two-Parametric Conditional Oscillation

Recall that h is a positive solution of (1.5) such that $h'(t) \neq 0$ for large t, $g = r\Phi(h')$ is its quasiderivative, R, G are given by (2.6), and $\hat{c}$ is given by (2.7). Recall also that the quasiderivative g is a solution of the reciprocal equation (2.12), denote by

$$\widetilde{G} := c^{1-q}g\Phi^{-1}(g') = -rh\Phi(h'), \qquad \widetilde{R} := c^{1-q}g^2\left|g'\right|^{q-2} = \frac{r^2|h'|^{2p-2}}{ch^{p-2}} \tag{4.1}$$

the "reciprocal" analogues of G and R, and define

$$\hat{r}(t) = \frac{1}{|h'(t)|^p\widetilde{R}(t)\left(\int^t \widetilde{R}^{-1}(s)\mathrm{d}s\right)^2}. \tag{4.2}$$

to Contents, page 1

Our main result reads as follows.

Theorem 4.1. *Suppose that conditions* (2.8) *hold. Further, suppose that*

$$\lim_{t\to\infty}\frac{\hat{r}(t)}{r(t)}=0, \tag{4.3}$$

and that there exist limits

$$\lim_{t\to\infty}\frac{r(t)|h'(t)|^p}{c(t)h^p(t)},\qquad \lim_{t\to\infty}\frac{\left(\hat{r}(t)\Phi\left(f'(t)\right)\right)'}{\hat{c}(t)\Phi\left(f(t)\right)}, \tag{4.4}$$

the second one being finite, where $f(t)=h(t)\left(\int^t R^{-1}(s)\mathrm{d}s\right)^{1/p}$. *If* $\mu-\lambda<1/2q$, *then* (1.9) *is nonoscillatory; if* $\mu-\lambda>1/2q$, *then it is oscillatory.*

Proof. First consider the case $\mu=0$ in (1.9), that is, we consider the equation

$$\left[\left(r(t)+\lambda\hat{r}(t)\right)\Phi\left(x'\right)\right]'+c(t)\Phi(x)=0. \tag{4.5}$$

The quantities $\widetilde{G}$ and $\widetilde{R}$ defined in (4.1) satisfy

$$\begin{aligned}\widetilde{G}&=-rh\Phi(h')=-G\,,\\ \widetilde{R}&=\frac{r^2|h'|^{2p-2}}{ch^{p-2}}=-\frac{h(r\Phi(h'))^2}{(r\Phi(h'))'},\end{aligned} \tag{4.6}$$

hence, integrating by parts,

$$\begin{aligned}\int^t\widetilde{R}^{-1}(s)\mathrm{d}s&=-\int^t\frac{1}{h(s)}\frac{[r(s)\Phi(h'(s))]'}{[r(s)\Phi(h'(s))]^2}\mathrm{d}s\\ &=\frac{1}{h(t)r(t)\Phi(h'(t))}+\int^t\frac{h'(s)}{h^2(s)}\frac{1}{r(s)\Phi(h'(s))}\mathrm{d}s\\ &=\frac{1}{G(t)}+\int^t R^{-1}(s)\mathrm{d}s.\end{aligned} \tag{4.7}$$

Consequently, conditions (2.8) imply that corresponding conditions for $\widetilde{G}$ and $\widetilde{R}$ also hold. This means, in view of Theorem 2.1 (applied to the reciprocal equation (2.12)), that the equation

$$\left(c^{1-q}(t)\Phi^{-1}(u')\right)'+\left[r^{1-q}(t)+\frac{\lambda}{|g(t)|^q\widetilde{R}(t)\left(\int^t\widetilde{R}^{-1}(s)\mathrm{d}s\right)^2}\right]\Phi^{-1}(u)=0 \tag{4.8}$$

is oscillatory for $\lambda>1/2p$ and nonoscillatory in the opposite case.

The reciprocal equation to (4.5) is the equation

$$\left(c^{1-q}(t)\Phi^{-1}(u')\right)' + (r(t) + \lambda\hat{r}(t))^{1-q}\Phi^{-1}(u) = 0. \tag{4.9}$$

Since (4.3) holds, we have

$$(r + \lambda\hat{r})^{1-q} = r^{1-q}\left(1 + \frac{\lambda\hat{r}}{r}\right)^{1-q} = r^{1-q}\left(1 + \frac{(1-q)\lambda\hat{r}}{r} + o\left(\frac{\hat{r}}{r}\right)\right), \tag{4.10}$$

as $t \to \infty$. Hence, we can rewrite (4.9) in the following form:

$$\left(c^{1-q}(t)\Phi^{-1}(u')\right)' + r^{1-q}(t)\left(1 + \frac{(1-q)\lambda\hat{r}(t)}{r(t)} + o\left(\frac{\hat{r}(t)}{r(t)}\right)\right)\Phi^{-1}(u) = 0. \tag{4.11}$$

Let $\lambda > -1/2q$ what is equivalent to $\lambda(1-q) < 1/2p$, then, in view of (4.3), there exists $\tilde{\lambda}$ such that $\lambda(1-q) < \tilde{\lambda} < 1/2p$, hence, for large t,

$$r^{1-q}\left(1 + \frac{\lambda(1-q)\hat{r}}{r} + o\left(\frac{\hat{r}}{r}\right)\right) < r^{1-q}\left(1 + \frac{\tilde{\lambda}\hat{r}}{r}\right) = r^{1-q} + \frac{\tilde{\lambda}}{|g|^q\tilde{R}\left(\int^t \tilde{R}^{-1}(s)\mathrm{d}s\right)^2}. \tag{4.12}$$

This means that the equation

$$\left(c^{1-q}(t)\Phi^{-1}(u')\right)' + \left[r^{1-q}(t) + \frac{\tilde{\lambda}}{|g(t)|^q\tilde{R}(t)\left(\int^t \tilde{R}^{-1}(s)\mathrm{d}s\right)^2}\right]\Phi^{-1}(u) = 0 \tag{4.13}$$

is a majorant of (4.9) and this majorant is nonoscillatory by Theorem 2.1 applied to (4.8). So (4.9) is also nonoscillatory, and hence (4.5) is nonoscillatory as well. The same argument implies oscillation of (4.5) if $\lambda < -1/2q$.

Now, we turn our attention to the general case $\mu \neq 0$. Let $f := h(\int^t R^{-1}(s)\mathrm{d}s)^{1/p}$, and consider the term

$$f\left[\ell(f) + \hat{\ell}(f)\right] \tag{4.14}$$

appearing in the modified Riccati equation (3.6), where the operators $\ell, \hat{\ell}$ are defined by (3.4). In order to use the asymptotic formula from Theorem 2.2, we write $f[\ell(f) + \hat{\ell}(f)] = A + B$, where

$$\begin{aligned} A &= f\left[(r\Phi(f'))' + \left(c + \frac{1}{2q}\hat{c}\right)\Phi(f)\right], \\ B &= f\left[\lambda(\hat{r}\Phi(f'))' + \left(\mu - \frac{1}{2q}\right)\hat{c}\Phi(f)\right]. \end{aligned} \tag{4.15}$$

Let $L \in \mathbb{R}$ be the second limit in (4.4), that is,

$$\left(\hat{r}\Phi(f')\right)' = L\hat{c}\Phi(f)(1 + o(1)) \quad \text{as } t \longrightarrow \infty. \tag{4.16}$$

The leading term in the expression A is const $G'G^{-2}(\int^t R^{-1}(s)ds)^{-1}$ by Theorem 2.2, while, concerning the asymptotics of B,

$$B = f\hat{c}\Phi(f)\left[L\lambda + \mu - \frac{1}{2q} + o(1)\right] = \frac{1}{R\left(\int^t R^{-1}(s)ds\right)}\left[L\lambda + \mu - \frac{1}{2q} + o(1)\right], \tag{4.17}$$

as $t \to \infty$. The existence of the first limit in (4.4) implies that there exists the limit

$$\begin{aligned} \lim_{t\to\infty} \frac{G'(t)G^{-2}(t)}{R^{-1}(t)} &= \lim_{t\to\infty} \frac{r(t)h^2(t)|h'(t)|^{p-2}\left(r(t)|h'(t)|^p - c(t)h^p(t)\right)}{(r(t)h(t)\Phi(h'(t)))^2} \\ &= 1 - \lim_{t\to\infty} \frac{c(t)h^p(t)}{r(t)|h'(t)|^p}. \end{aligned} \tag{4.18}$$

The limit in (4.18) must be 0, which follows from the l'Hospital rule and the fact that the integral of R^{-1} is divergent, while the integral of $G'G^{-2}$ is convergent by the second assumption in (2.8). This means that the term B dominates A; hence, $A(t) + B(t) > 0$ for large t if $L\lambda + \mu - 1/2q > 0$ and $A(t) + B(t) < 0$ for large t if $L\lambda + \mu - 1/2q < 0$.

Now, it remains to prove that these inequalities imply (non)oscillation of (1.9) and that $L = -1$.

To prove the nonoscillation, let $L\lambda + \mu - 1/2q < 0$, that is, $A(t) + B(t) < 0$ for large t, and let $\mathcal{G}$ be defined by (3.5). By Lemma 3.3(i) $v = 0$ is a solution of the inequality

$$v' + A(t) + B(t) + (p-1)(r(t) + \lambda\hat{r}(t))^{1-q} f^{-q}(t)\mathcal{G}(t, v) \le 0, \tag{4.19}$$

for large t, and by identity (3.3) in Lemma 3.2 we obtain that $w = (r + \lambda\hat{r})\Phi(f'/f)$ satisfies the Riccati inequality (3.2), that is, (1.9) is nonoscillatory by Lemma 3.1(iii).

To prove the oscillation, let $L\lambda + \mu - 1/2q > 0$, that is, $A(t) + B(t) > 0$ for large t. Observe that for $t \to \infty$

$$\int^t f^p(s)\hat{c}(s)ds = \int^t \frac{1}{R(s)\left(\int^t R^{-1}(\tau)d\tau\right)}ds = \log\left(\int^t R^{-1}(s)ds\right) \longrightarrow \infty, \tag{4.20}$$

and hence $\int^\infty B(t)dt = \infty$, which consequently means that $\int^\infty (A(t) + B(t))dt = \infty$. Here, we have used the fact that the integral of the leading term in A and also integrals of other terms in the asymptotic formula of Theorem 2.2 are convergent, see [7, page 161]. Suppose, on the contrary, that (1.9) is nonoscillatory. Then by Lemma 3.1, there exists a solution w of the associated Riccati equation (3.1) for large t and, by Lemma 3.2, the function $v = f^p(w - w_f)$,

to Contents, page 1

where $w_f = (r + \lambda\hat{r})\Phi(f'/f)$, is a solution of the modified Riccati equation (3.6) for large t. Integrating (3.6), we get

$$\begin{aligned} v(T) - v(t) = {} & \int_T^t (A(s) + B(s))\,\mathrm{d}s \\ & + (p-1)\int_T^t (r(s) + \lambda\hat{r}(s))^{1-q} f^{-q}(s)\mathcal{G}(s, v(s))\,\mathrm{d}s. \end{aligned} \tag{4.21}$$

Now, we use Theorem 3.5. In view of (2.8) and (4.3), we have for $t \to \infty$,

$$\begin{aligned} |\Omega(t)| &= (r(t) + \lambda\hat{r}(t)) f(t) \left|\Phi(f'(t))\right| \\ &= r(t)(1 + o(1))h(t)\left(\int^t R^{-1}(s)\mathrm{d}s\right)^{1/p} \left|\Phi(h'(t))\right| \left(\int^t R^{-1}(s)\mathrm{d}s\right)^{(p-1)/p} \\ &\quad \times \left(1 + \frac{1}{pG(t)\left(\int^t R^{-1}(s)\mathrm{d}s\right)}\right)^{p-1} \\ &= |G(t)|\left(\int^t R^{-1}(s)\mathrm{d}s\right)(1 + o(1)) \longrightarrow \infty, \\ \mathcal{R}(t) &= (r(t) + \lambda\hat{r}(t)) f^2(t) \left|f'(t)\right|^{p-2} \\ &= r(t)(1 + o(1))h^2(t)\left(\int^t R^{-1}(s)\mathrm{d}s\right)^{2/p} \left|h'(t)\right|^{p-2} \left(\int^t R^{-1}(s)\mathrm{d}s\right)^{(p-2)/p} \\ &\quad \times \left(1 + \frac{1}{pG(t)\left(\int^t R^{-1}(s)\mathrm{d}s\right)}\right)^{p-2} \\ &= R(t)\left(\int^t R^{-1}(s)\mathrm{d}s\right)(1 + o(1)), \end{aligned} \tag{4.22}$$

and hence

$$\int^t \frac{\mathrm{d}s}{\mathcal{R}(s)} \longrightarrow \infty \quad \text{as } t \longrightarrow \infty. \tag{4.23}$$

Consequently, $v(t) \geq 0$ by Theorem 3.5. This means that the left-hand side in (4.21) is bounded above as $t \to \infty$, while the right-hand side tends to ∞ which yields the required contradiction proving that (1.9) is oscillatory if $L\lambda + \mu > 1/2q$.

Finally, consider again the case $\mu = 0$. In that case, we proved in the first part of the proof that (1.9) is oscillatory or nonoscillatory depending on whether $\lambda < -1/2q$ or $\lambda > -1/2q$. This shows that the second limit in (4.4) must be -1. □

to Contents, page 1

Remark 4.2. (i) From the proof of Theorem 4.1, it follows that if the first limit in (4.4) exists, then conditions (2.8) imply that this limit is 1, and the assumptions of the theorem imply that if the second limit in (4.4) exists and is finite, then it is -1.

(ii) Theorem 4.1 can be applied to the Euler equation (1.7), and one can obtain the same result for (1.8) as in [4, Corollary 3]. Indeed, in this case, we have $h(t) = t^{(p-1)/p}$, $r = 1$, $c(t) = \gamma_p t^{-p}$, where $\gamma_p = ((p-1)/p)^p$ and by a direct computation

$$G(t) = \left(\frac{p-1}{p}\right)^{p-1}, \qquad R(t) = \widetilde{R}(t) = \left(\frac{p-1}{p}\right)^{p-2} t, \tag{4.24}$$

hence,

$$\begin{aligned}
\widehat{c}(t) &= \left[\left(t^{(p-1)/p}\right)^p \left(\frac{p-1}{p}\right)^{p-2} t\left[\left(\frac{p}{p-1}\right)^{p-2}\log t\right]^2\right]^{-1} = \left(\frac{p}{p-1}\right)^{2-p} t^{-p}\log^{-2} t,\\
\widehat{r}(t) &= \left[\left(\frac{p-1}{p}t^{-1/p}\right)^p \left(\frac{p-1}{p}\right)^{p-2} t\left[\left(\frac{p}{p-1}\right)^{p-2}\log t\right]^2\right]^{-1} = \left(\frac{p}{p-1}\right)^{2}\log^{-2} t,
\end{aligned}\tag{4.25}$$

which mean that conditions (2.8) and (4.3) are satisfied. Concerning the limits in (4.4), we have

$$r\left|h'(t)\right|^p = \left(\frac{p-1}{p}\right)^p t^{-1} = c(t)h^p(t), \tag{4.26}$$

that is, the first limit in (4.4) is 1. Next,

$$f(t) = \left(\frac{p}{p-1}\right)^{(p-2)/p} t^{(p-1)/p}\log^{1/p} t, \tag{4.27}$$

and consequently,

$$\begin{aligned}
\widehat{c}(t)\Phi(f(t)) &= \left(\frac{p}{p-1}\right)^{2-p} t^{-p}\log^{-2} t\left[\left(\frac{p}{p-1}\right)^{(p-2)/p} t^{(p-1)/p}\log^{1/p} t\right]^{p-1}\\
&= \left(\frac{p}{p-1}\right)^{-1+2/p} t^{-2+1/p}\log^{-1-1/p} t,\\
f'(t) &= \left(\frac{p}{p-1}\right)^{(p-2)/p}\left[\frac{p-1}{p}t^{-1/p}\log^{1/p} t + \frac{1}{p}t^{-1/p}\log^{1/p-1} t\right]\\
&= \left(\frac{p-1}{p}\right)^{2/p} t^{-1/p}\log^{1/p} t\left[1 + \frac{1}{p-1}\log^{-1} t\right].
\end{aligned}\tag{4.28}$$

to Contents, page 1

Using this formula,

$$\hat{r}(t)\Phi(f'(t)) = \left(\frac{p}{p-1}\right)^{2/p} t^{-1+1/p}\log^{-1-1/p}t\left[1+\log^{-1}t+\frac{p-2}{2}\log^{-2}t+O\left(\log^{-2}t\right)\right], \tag{4.29}$$

and hence,

$$\begin{aligned}(\hat{r}(t)\Phi(f'(t)))' &= \left(\frac{p}{p-1}\right)^{2/p}\left[-\frac{p-1}{p}t^{-2+1/p}\log^{-1-1/p}t\left(1+O\left(\log^{-1}t\right)\right)\right.\\ &\qquad -\frac{p+1}{p}t^{-2+1/p}\log^{-2-1/p}t\left(1+O\left(\log^{-1}t\right)\right)\\ &\qquad \left.+t^{-2+1/p}\log^{-1-1/p}tO\left(\log^{-2}t\right)\right]\\ &= -\left(\frac{p}{p-1}\right)^{2/p-1} t^{-2+1/p}\log^{-1-1/p}t\left(1+O\left(\log^{-1}t\right)\right),\end{aligned} \tag{4.30}$$

as $t \to \infty$. This means that the second limit in (4.4) is -1. According to Theorem 4.1, we obtain that the equation

$$\left[\left(1+\lambda\left(\frac{p}{p-1}\right)^{2}\frac{1}{\log^2 t}\right)\Phi(x')\right]' + \left[\frac{\gamma_p}{t^p}+\mu\left(\frac{p}{p-1}\right)^{2-p}\frac{1}{t^p\log^2 t}\right]\Phi(x) = 0 \tag{4.31}$$

is nonoscillatory if $\mu-\lambda < 1/2q$ and oscillatory if $\mu-\lambda > 1/2q$. If we denote $\tilde{\lambda} = \lambda(p/(p-1))^2$ and $\tilde{\mu} = \mu(p/(p-1))^{2-p}$, we see that (1.8) (with $\tilde{\lambda}$, $\tilde{\mu}$ instead of λ, μ, resp.) is nonoscillatory if $\tilde{\mu}-\gamma_p\tilde{\lambda} < (1/2)((p-1)/p)^{p-1}$, and it is oscillatory if $\tilde{\mu}-\gamma_p\tilde{\lambda} > (1/2)((p-1)/p)^{p-1}$, that is, we have the statement from [4].

(iii) In [3], it is proved that (1.8) is nonoscillatory also in the limiting case $\mu-\gamma_p\lambda = \mu_p$. We conjecture that we have also the same situation in the general case, that is, (1.9) is nonoscillatory also in the case $\mu-\lambda = 1/2q$.

Acknowledgments

This research was supported by Grants nos. P201/11/0768 and P201/10/1032 of the Czech Science Foundation and the Research Project no. MSM0021622409 of the Ministry of Education of the Czech Republic.

References

[1] E. Hille, "Non-oscillation theorems," *Transactions of the American Mathematical Society*, vol. 64, pp. 234–252, 1948.
[2] T. Kusano, Y. Naito, and T. Tanigawa, "Strong oscillation and nonoscillation of quasilinear differential equations of second order," *Differential Equations and Dynamical Systems*, vol. 2, no. 1, pp. 1–10, 1994.
[3] O. Došlý and H. Haladová, "Half-linear Euler differential equations in the critical case," to appear in *Tatra Mountains Mathematical Publications*.

to Contents, page 1

[4] O. Došlý and S. Fišnarová, "Half-linear oscillation criteria: Perturbation in term involving derivative," *Nonlinear Analysis, Theory, Methods and Applications*, vol. 73, no. 12, pp. 3756–3766, 2010.
[5] P. Hartman, *Ordinary Differential Equations*, Birkhäuser, Boston, Mass, USA, 2nd edition, 1982.
[6] F. Gesztesy and M. Ünal, "Perturbative oscillation criteria and Hardy-type inequalities," *Mathematische Nachrichten*, vol. 189, pp. 121–144, 1998.
[7] O. Došlý and M. Ünal, "Conditionally oscillatory half-linear differential equations," *Acta Mathematica Hungarica*, vol. 120, no. 1-2, pp. 147–163, 2008.
[8] O. Došlý and P. Řehák, *Half-Linear Differential Equations*, vol. 202 of *North-Holland Mathematics Studies*, Elsevier, Amsterdam, The Netherlands, 2005.
[9] O. Došlý and S. Fišnarová, "Variational technique and principal solution in half-linear oscillation criteria," *Applied Mathematics and Computation*, vol. 217, no. 12, pp. 5385–5391, 2011.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 543035, 13 pages
doi:10.1155/2011/543035

Research Article

Uniqueness of Positive Solutions for a Class of Fourth-Order Boundary Value Problems

J. Caballero, J. Harjani, and K. Sadarangani

Departamento de Matemáticas, Universidad de Las Palmas de Gran Canaria, Campus de Tafira Baja, 35017 Las Palmas de Gran Canaria, Spain

Correspondence should be addressed to K. Sadarangani, ksadaran@dma.ulpgc.es

Received 31 December 2010; Accepted 23 March 2011

Academic Editor: Yuri V. Rogovchenko

The purpose of this paper is to investigate the existence and uniqueness of positive solutions for the following fourth-order boundary value problem: $y^{(4)}(t) = f(t, y(t))$, $t \in [0,1]$, $y(0) = y(1) = y'(0) = y'(1) = 0$. Moreover, under certain assumptions, we will prove that the above boundary value problem has a unique symmetric positive solution. Finally, we present some examples and we compare our results with the ones obtained in recent papers. Our analysis relies on a fixed point theorem in partially ordered metric spaces.

1. Introduction

The purpose of this paper is to consider the existence and uniqueness of positive solutions for the following fourth-order two-point boundary value problem:

$$\begin{aligned} &y^{(4)}(t) = f(t, y(t)), \quad t \in [0,1], \\ &y(0) = y(1) = y'(0) = y'(1) = 0, \end{aligned} \tag{1.1}$$

which describes the bending of an elastic beam clamped at both endpoints.

There have been extensive studies on fourth-order boundary value problems with diverse boundary conditions. Some of the main tools of nonlinear analysis devoted to the study of this type of problems are, among others, lower and upper solutions [1–4], monotone iterative technique [5–7], Krasnoselskii fixed point theorem [8], fixed point index [9–11], Leray-Schauder degree [12, 13], and bifurcation theory [14–16].

to Contents, page 1

2. Background

In this section, we present some basic facts which are necessary for our results.

In our study, we will use a fixed point theorem in partially ordered metric spaces which appears in [17].

Let $\mathcal{M}$ denote the class of those functions $\beta : [0,\infty) \rightarrow [0,1)$ satisfying the condition

$$\beta(t_n) \longrightarrow 1 \quad \text{implies } t_n \rightarrow 0. \tag{2.1}$$

Now, we recall the above mentioned fixed point theorem.

Theorem 2.1 (see 1, Theorem 2.1). *Let $(X,\leq)$ be a partially ordered set and suppose that there exists a metric d in X such that (X,d) is a complete metric space. Let $T : X \rightarrow X$ be a nondecreasing mapping such that there exists an element $x_0 \in X$ with $x_0 \leq Tx_0$. Suppose that there exists $\beta \in \mathcal{M}$ such that*

$$d(Tx,Ty) \leq \beta(d(x,y)) \cdot d(x,y), \quad \text{for any } x,y \in X \text{ with } x \geq y. \tag{2.2}$$

Assume that either T is continuous or X is such that

$$\text{if } (x_n) \text{ is a nondecreasing sequence in } X \text{ such that } x_n \longrightarrow x, \text{ then } x_n \leq x \text{ for all } n \in \mathbb{N}. \tag{2.3}$$

Besides, suppose that

$$\text{for each } x,y \in X, \text{ there exists } z \in X \text{ which is comparable to } x \text{ and } y. \tag{2.4}$$

Then T has a unique fixed point.

In our considerations, we will work with a subset of the classical Banach space $C[0,1]$. This space will be considered with the standard metric

$$d(x,y) = \sup_{0\leq t\leq 1} |x(t) - y(t)|. \tag{2.5}$$

This space can be equipped with a partial order given by

$$x,y \in C[0,1], \quad x \leq y \Longleftrightarrow x(t) \leq y(t), \quad \text{for } t \in [0,1]. \tag{2.6}$$

In [18], it is proved that $(C[0,1],\leq)$ with the above mentioned metric satisfies condition (2.3) of Theorem 2.1. Moreover, for $x,y \in C[0,1]$, as the function $\max(x,y) \in C[0,1]$, $(C[0,1],\leq)$ satisfies condition (2.4).

On the other hand, the boundary value problem (1.1) can be rewritten as the integral equation (see, e.g., [19])

$$y(t) = \int_0^1 G(t,s) f(s,u(s)) \mathrm{d}s, \quad \text{for } t \in [0,1], \tag{2.7}$$

to Contents, page 1

where $G(t,s)$ is the Green's function given by

$$G(t,s) = \frac{1}{6}\begin{cases} t^2(1-s)^2[(s-t)+2(1-t)s], & 0 \leq t \leq s \leq 1, \\ s^2(1-t)^2[(t-s)+2(1-s)t], & 0 \leq s \leq t \leq 1. \end{cases} \tag{2.8}$$

Note that $G(t,s)$ satisfies the following properties:

(i) $G(t,s)$ is a continuous function on $[0,1] \times [0,1]$,

(ii) $G(0,s) = G(1,s) = 0$, for $s \in [0,1]$,

(iii) $G(t,s) \geq 0$, for $t,s \in [0,1]$.

3. Main Results

Our starting point in this section is to present the class of functions $\mathcal{A}$ which we use later. By $\mathcal{A}$ we denote the class of functions $\phi : [0,\infty) \to [0,\infty)$ satisfying the following conditions:

(i) ϕ is nondecreasing,

(ii) for any $x > 0$, $\phi(x) < x$,

(iii) $\beta(x) = \phi(x)/x \in \mathcal{M}$.

Examples of functions in $\mathcal{A}$ are $\phi(x) = \mu x$ with $0 \leq \mu < 1$, $\phi(x) = x/(1+x)$ and $\phi(x) = \ln(1+x)$. In the sequel, we formulate our main result.

Theorem 3.1. *Consider problem* (1.1) *assuming the following hypotheses:*

(a) $f : [0,1] \times [0,\infty) \to [0,\infty)$ *is continuous,*

(b) $f(t,y)$ *is nondecreasing with respect to the second variable, for each* $t \in [0,1]$,

(c) *suppose that there exists* $0 < \alpha \leq 384$, *such that, for* $x,y \in [0,\infty)$ *with* $y \geq x$,

$$f(t,y) - f(t,x) \leq \alpha\phi(y-x), \quad \text{with } \phi \in \mathcal{A}. \tag{3.1}$$

Then, problem (1.1) *has a unique nonnegative solution.*

Proof. Consider the cone

$$P = \{x \in C[0,1] : x \geq 0\}. \tag{3.2}$$

Obviously, (P,d) with $d(x,y) = \sup\{|x(t) - y(t)| : t \in [0,1]\}$ is a complete metric space satisfying condition (2.3) and condition (2.4) of Theorem 2.1.

to Contents, page 1

Consider the operator defined by

$$(Tx)(t) = \int_0^1 G(t,s) f(s, x(s))\mathrm{d}s, \quad \text{for } x \in P, \tag{3.3}$$

where $G(t,s)$ is the Green's function defined in Section 2.

It is clear that T applies the cone P into itself since $f(t,x)$ and $G(t,s)$ are nonnegative continuous functions.

Now, we check that assumptions in Theorems 2.1 are satisfied.

Firstly, the operator T is nondecreasing.

Indeed, since f is nondecreasing with respect to the second variable, for $u, v \in P$, $u \geq v$ and $t \in [0,1]$, we have

$$\begin{aligned}(Tu)(t) &= \int_0^1 G(t,s) f(s,u(s))\mathrm{d}s \\ &\geq \int_0^1 G(t,s) f(s,v(s))\mathrm{d}s \\ &= (Tv)(t).\end{aligned} \tag{3.4}$$

On the other hand, a straightforward calculation gives us

$$\begin{gathered}\int_0^1 G(t,s)\mathrm{d}s = \int_0^t G(t,s)\mathrm{d}s + \int_t^1 G(t,s)\mathrm{d}s = \frac{t^2}{24} - \frac{t^3}{12} + \frac{t^4}{24}, \\ \max_{0\leq t\leq 1} \int_0^1 G(t,s)\mathrm{d}s = \max_{0\leq t\leq 1}\left(\frac{t^2}{24} - \frac{t^3}{12} + \frac{t^4}{24}\right) = \frac{1}{384}.\end{gathered} \tag{3.5}$$

Taking into account this fact and our hypotheses, for $u, v \in P$ and $u > v$, we can obtain the following estimate:

$$\begin{aligned}d(Tu,Tv) &= \sup_{0\leq t\leq 1}\Big|(Tu)(t) - (Tv)(t)\Big| \\ &= \sup_{0\leq t\leq 1}((Tu)(t) - (Tv)(t)) \\ &= \sup_{0\leq t\leq 1}\int_0^1 G(t,s)\big(f(s,u(s)) - f(s,v(s))\big)\mathrm{d}s \\ &\leq \sup_{0\leq t\leq 1}\int_0^1 G(t,s)\alpha\phi(u(s) - v(s))\mathrm{d}s \\ &\leq \sup_{0\leq t\leq 1}\int_0^1 G(t,s)\alpha\phi(d(u,v))\mathrm{d}s\end{aligned}$$

to Contents, page 1

$$
\begin{aligned}
&= \alpha\phi(d(u,v)) \sup_{0\le t\le 1}\int_0^1 G(t,s)\mathrm{d}s \\
&= \alpha\phi(d(u,v)) \cdot \frac{1}{384} \\
&\le \phi(d(u,v)) \\
&= \frac{\phi(d(u,v))}{d(u,v)} \cdot d(u,v).
\end{aligned} \tag{3.6}
$$

This gives us, for $u, v \in P$ and $u > v$,

$$d(Tu, Tv) \le \beta(d(u,v)) \cdot d(u,v), \tag{3.7}$$

where $\beta(x) = \phi(x)/x \in \mathcal{M}$.

Obviously, the last inequality is satisfied for $u = v$.

Therefore, the contractive condition appearing in Theorem 2.1 is satisfied for $u \ge v$. Besides, as f and G are nonnegative functions,

$$T0 = \int_0^1 G(t,s) f(s,0)\mathrm{d}s \ge 0. \tag{3.8}$$

Finally, Theorem 2.1 tells us that T has a unique fixed point in P, and this means that problem (1.1) has a unique nonnegative solution.

This finishes the proof. □

Now, we present a sufficient condition for the existence and uniqueness of positive solutions for our problem (1.1) (positive solution means $x(t) > 0$, for $t \in (0,1)$). The proof of the following theorem is similar to the proof of Theorem 3.6 of [8]. We present a proof for completeness.

Theorem 3.2. *Under assumptions of Theorem 3.1 and suppose that $f(t_0, 0) \ne 0$ for certain $t_0 \in [0,1]$, problem* (1.1) *has a unique positive solution.*

Proof. Consider the nonnegative solution $x(t)$ given by Theorem 3.1 of problem (1.1).

Notice that this solution satisfies

$$x(t) = \int_0^1 G(t,s) f(s, x(s))\mathrm{d}s. \tag{3.9}$$

Now, we will prove that x is a positive solution.

to Contents, page 1

In contrary case, suppose that there exists $0 < t^* < 1$ such that $x(t^*) = 0$ and, consequently,

$$x(t^*) = \int_0^1 G(t^*, s) f(s, x(s)) \mathrm{d}s = 0. \tag{3.10}$$

Since $x \geq 0$, f is nondecreasing with respect to the second variable and $G(t,s) \geq 0$, we have

$$0 = x(t^*) = \int_0^1 G(t^*, s) f(s, x(\mathrm{s})) \mathrm{d}s \geq \int_0^1 G(t^*, s) f(s, 0) \mathrm{d}s \geq 0, \tag{3.11}$$

and this gives us

$$\int_0^1 G(t^*, s) f(s, 0) \mathrm{d}s = 0. \tag{3.12}$$

This fact and the nonnegative character of $G(t,s)$ and $f(t,x)$ imply

$$G(t^*, s) \cdot f(s, 0) = 0 \quad \mathrm{a} \cdot \mathrm{e} \ (s). \tag{3.13}$$

As $G(t^*, s) \neq 0 \mathrm{a{\cdot}e}\ (s)$, because $G(t^*, s)$ is given by a polynomial, we obtain

$$f(s, 0) = 0 \quad \mathrm{a} \cdot \mathrm{e} \ (s). \tag{3.14}$$

On the other hand, as $f(t_0, 0) \neq 0$ for certain $t_0 \in [0,1]$ and $f(t_0, x) \geq 0$, we have that $f(t_0, 0) > 0$.

The continuity of f gives us the existence of a set $A \subset [0,1]$ with $t_0 \in A$ and $\mu(A) > 0$, where μ is the Lebesgue measure, satisfying that $f(t, 0) > 0$ for any $t \in A$. This contradicts (3.14).

Therefore, $x(t) > 0$ for $t \in (0,1)$.

This finishes the proof. □

Now, we present an example which illustrates our results.

Example 3.3. Consider the nonlinear fourth-order two-point boundary value problem

$$\begin{gathered} y^{(4)}(t) = c + \lambda \arctan\big(y(t)\big), \quad t \in (0,1),\ c, \lambda > 0, \\ y(0) = y(1) = y'(0) = y'(1) = 0. \end{gathered} \tag{3.15}$$

In this case, $f(t, y) = c + \lambda \arctan y$. It is easily seen that $f(t, y)$ satisfies (a) and (b) of Theorem 3.1.

to Contents, page 1

In order to prove that $f(t, y)$ satisfies (c) of Theorem 3.1, previously, we will prove that the function $\phi : [0, \infty) \to [0, \infty)$, defined by $\phi(x) = \arctan x$, satisfies

$$\phi(u) - \phi(v) \leq \phi(u - v) \quad \text{for } u \geq v. \tag{3.16}$$

In fact, put $\phi(u) = \arctan u = \alpha$ and $\phi(v) = \arctan v = \beta$ (notice that, as $u \geq v$ and ϕ is nondecreasing, $\alpha \geq \beta$). Then, from

$$\tan(\alpha - \beta) = \frac{\tan\alpha - \tan\beta}{1 + \tan\alpha \cdot \tan\beta}, \tag{3.17}$$

as $\alpha, \beta \in [0, \pi/2)$, then $\tan\alpha, \tan\beta \in [0, \infty)$, we obtain

$$\tan(\alpha - \beta) \leq \tan\alpha - \tan\beta. \tag{3.18}$$

Applying ϕ to this inequality and taking into account the nondecreasing character of ϕ, we have

$$\alpha - \beta \leq \arctan(\tan\alpha - \tan\beta) \tag{3.19}$$

or, equivalently,

$$\phi(u) - \phi(v) = \arctan u - \arctan v \leq \arctan(u - v) = \phi(u - v). \tag{3.20}$$

This proves our claim.

In the sequel, we prove that $f(t, y)$ satisfies assumption (c) of Theorem 3.1.

In fact, for $y \geq x$ and $t \in [0, 1]$, we can obtain

$$\begin{aligned} f(t, y) - f(t, x) &= \lambda(\arctan y - \arctan x) \\ &\leq \lambda \arctan(y - x). \end{aligned} \tag{3.21}$$

Now, we will prove that $\phi(x) = \arctan x$ belongs to $\mathcal{A}$. In fact, obviously ϕ takes $[0, \infty)$ into itself and, as $\phi'(x) = 1/(1+x^2)$, ϕ is nondecreasing. Besides, as the derivative of $\psi(x) = x - \phi(x)$ is $\psi'(x) = 1 - 1/(1 + x^2) > 0$ for $x > 0$, ψ is strictly increasing, and, consequently, $\phi(x) < x$ for $x > 0$ (notice that $\psi(0) = 0$). Notice that if $\beta(x) = \phi(x)/x = \arctan x/x$ and $\beta(t_n) \to 1$, then (t_n) is a bounded sequence because, in contrary case, $t_n \to \infty$ and, thus, $\beta(t_n) \to 0$. Suppose that $t_n \nrightarrow 0$. Then, we can find $\epsilon > 0$ such that, for each $n \in \mathbb{N}$, there exists $p_n \geq n$ with $t_{p_n} \geq \epsilon$. The bounded character of (t_n) gives us the existence of a subsequence (t_{k_n}) of (t_{p_n}) with (t_{k_n}) convergent. Suppose that $t_{k_n} \to a$. From $\beta(t_n) \to 1$, we obtain $\arctan t_{k_n}/t_{k_n} \to \arctan a/a = 1$ and, as the unique solution of $\arctan x = x$ is $x_0 = 0$, we obtain $a = 0$. Thus, $t_{k_n} \to 0$, and this contradicts the fact that $t_{k_n} \geq \epsilon$ for any $n \in \mathbb{N}$. Therefore, $t_n \to 0$. This proves that $f(t, y)$ satisfies assumption (c) of Theorem 3.1. Finally, as $f(t, 0) = c > 0$, Problem (3.15) has a unique positive solution for $0 < \lambda \leq 384$ by Theorems 3.1 and 3.2.

to Contents, page 1

Remark 3.4. In Theorem 3.2, the condition $f(t_0,0) \neq 0$ for certain $t_0 \in [0,1]$ seems to be a strong condition in order to obtain a positive solution for Problem (1.1), but when the solution is unique, we will see that this condition is very adjusted one. More precisely, under assumption that Problem (1.1) has a unique nonnegative solution $x(t)$, then

$$f(t,0) = 0 \quad \text{for } t \in [0,1] \text{ iff } x(t) \equiv 0. \tag{3.22}$$

In fact, if $f(t,0) = 0$ for $t \in [0,1]$, then it is easily seen that the zero function satisfies Problem (1.1) and the uniqueness of solution gives us $x(t) \equiv 0$.

The other implication is obvious since if the zero function is solution of Problem (1.1), then $0 = f(t,0)$ for any $t \in [0,1]$.

Remark 3.5. Notice that assumptions in Theorem 3.1 are invariant by continuous perturbations. More precisely, if $f(t,0) = 0$ for any $t \in [0,1]$ and f satisfies (a), (b), and (c) of Theorem 3.1, then $g(t,x) = a(t) + f(t,x)$, with $a : [0,1] \to [0,\infty)$ continuous and $a \neq 0$, satisfies assumptions of Theorem 3.2, and this means that the following boundary value problem

$$\begin{gathered} y^{(4)}(t) = g(t, y(t)), \quad t \in [0,1], \\ y(0) = y(1) = y'(0) = y'(1) = 0, \end{gathered} \tag{3.23}$$

has a unique positive solution.

4. Some Remarks

In this section, we compare our results with the ones obtained in recent papers. Recently, in [19], the authors present as main result the following theorem.

Theorem 4.1 (Theorem 3.1 of [19]). *Suppose that*

($H1$) $f : [0,1] \times [0,\infty) \to [0,\infty)$ *is continuous,*

($H2$) $f(t,y)$ *is nondecreasing in* y, *for each* $t \in [0,1]$,

($H3$) $f(t,y) = f(1-t,y)$ *for each* $(t,y) \in [0,1] \times [0,\infty)$.

Moreover, suppose that there exist positive numbers $a > b$ *such that*

$$\max_{0 \le t \le 1} f(t,a) \le a \cdot A, \qquad \min_{1/4 \le t \le 3/4} f\left(t, \frac{b}{16}\right) \ge b \cdot B, \tag{4.1}$$

where

$$A = \left(\max_{0 \le t \le 1} \int_0^1 G(t,s)\,\mathrm{d}s\right)^{-1}, \qquad B = \left(\max_{0 \le t \le 1} \int_{1/4}^{3/4} G(t,s)\,\mathrm{d}s\right)^{-1}, \tag{4.2}$$

to Contents, page 1

with $G(t,s)$ *being the Green's function defined in Section 2.Then, Problem* (1.1) *has at least one symmetric positive solution* $y^* \in \mathcal{C}[0,1]$ *such that* $b \le \|y^*\| \le a$ *and, moreover,* $y^* = \lim_{k\to\infty} T^k y_0$ *in the uniform norm, where* T *is the operator defined by*

$$(Tx)(t) = \int_0^1 G(t,s) f(s,x(s))\mathrm{d}s, \quad \textit{for } x \in \mathcal{C}[0,1] \tag{4.3}$$

and y_0 *is the function given by* $y_0(t) = b \cdot q(t)$, *for* $t \in [0,1]$, *with* $q(t) = \min(t^2, (1-t)^2)$, *for* $t \in [0,1]$ *(symmetric solution means a solution* $y(t)$ *satisfying* $y(t) = y(1-t)$, *for* $t \in [0,1]$*).*

In what follows, we present a parallel result to Theorem 3.2 where we obtain uniqueness of a symmetric positive solution of Problem (1.1).

Theorem 4.2. *Adding assumption* $(H3)$ *of Theorem 4.1 to the hypotheses of Theorem 3.2, one obtains a unique symmetric positive solution of Problem* (1.1).

Proof. As in the proof of Theorem 3.1, instead of P, we consider the following set K

$$K = \{x \in \mathcal{C}[0,1] : x \ge 0 \ \text{ and } x \text{ is symmetric}\}. \tag{4.4}$$

It is easily seen that K is a closed subset of $\mathcal{C}[0,1]$. Thus, (K,d), where d is the induced metric given by

$$d(x,y) = \sup_{0\le t\le 1} |x(t) - y(t)|, \quad \text{for } x,y \in K, \tag{4.5}$$

is a complete metric space.

Moreover, K with the induced order by $(\mathcal{C}[0,1], \le)$ satisfies condition (2.3) of Theorem 2.1, and it is easily proved that the function $\max(x,y) \in K$, for $x,y \in K$ and, consequently, $(K,\le)$, satisfies condition (2.4) of Theorem 2.1.

Now, as in Theorem 2.1, we consider the operator defined by

$$(Tx)(t) = \int_0^1 G(t,s) f(s,x(s))\mathrm{d}s, \quad \text{for } x \in K. \tag{4.6}$$

In the sequel, we prove that, under our assumptions, T applies K into itself.

In fact, suppose that x is symmetric, then for $t \in [0,1]$, we have

$$(Tx)(1-t) = \int_0^1 G(1-t,s) f(s,x(s))\mathrm{d}s. \tag{4.7}$$

to Contents, page 1

Making the change of variables $s = 1 - u$, we obtain

$$(Tx)(1-t) = -\int_1^0 G(1-t, 1-u) f(1-u, x(1-u)) \mathrm{d}u$$
$$= \int_0^1 G(1-t, 1-u) f(1-u, x(1-u)) \mathrm{d}u. \tag{4.8}$$

Now, it is easily seen that $G(t,s) = G(1-t, 1-s)$ for $t, s \in [0,1]$ and taking into account assumption $(H3)$ of Theorem 4.1 and the symmetric character of x, we have

$$\begin{aligned}(Tx)(1-t) &= \int_0^1 G(t,u) f(u, x(1-u)) \mathrm{d}u \\ &= \int_0^1 G(t,u) f(u, x(u)) \mathrm{d}u \\ &= (Tx)(t).\end{aligned} \tag{4.9}$$

The rest of the proof follows the lines of Theorems 3.1 and 3.2.

This finishes the proof. □

Now, we present an example which illustrates Theorem 4.2.

Example 4.3. Consider the following problem

$$\begin{gathered}y^{(4)}(t) = c + \lambda \sin(\pi t) \arctan(y(t)), \quad t \in (0,1),\ c, \lambda > 0, \\ y(0) = y(1) = y'(0) = y'(1) = 0.\end{gathered} \tag{4.10}$$

In this case, $f(t,y) = c + \lambda \sin(\pi t) \arctan y$. It is easily checked that $f(t,y)$ satisfies (a) and (b) of Theorem 3.1 and $f(t,y) = f(1-t, y)$, for $(t,y) \in [0,1] \times [0,\infty)$.

On the other hand, taking into account Example 3.3, we can obtain, for $y \geq x$ and $t \in [0,1]$,

$$\begin{aligned}f(t,y) - f(t,x) &= \lambda \sin \pi t [\arctan y - \arctan x] \\ &\leq \lambda \sin \pi t [\arctan(y-x)] \\ &\leq \lambda \arctan(y-x).\end{aligned} \tag{4.11}$$

Finally, as it is proved in Example 3.3, $\phi(x) = \arctan x$ belongs to $\mathcal{A}$. Therefore, Theorem 4.2 tells us that Problem (4.10) has a unique symmetric positive solution for $0 < \lambda \leq 384$. In what follows, we prove that Problem (4.10) can be treated using Theorem 4.1. In fact, in this case, $f(t,y) = c + \lambda \sin(\pi t) \arctan y$. Moreover, $A = 384$ (see proof of Theorem 3.1);

to Contents, page 1

it can be proved that $B = 531.61$. As we have seen in Example 4.3, $f(t,y)$ satisfies assumptions $(H1)$, $(H2)$, and $(H3)$ of Theorem 4.1. Moreover,

$$\begin{gathered}\max_{0\le t\le 1} f(t,a) = f\left(\frac{1}{2},a\right) = c+\lambda \arctan a,\\ \min_{1/4\le t\le 3/4} f\left(t,\frac{b}{16}\right) = f\left(\frac{1}{4},\frac{b}{16}\right) = c+\lambda\sin\frac{\pi}{4}\arctan\left(\frac{b}{16}\right) = c+\lambda\frac{\sqrt{2}}{2}\arctan\left(\frac{b}{16}\right).\end{gathered} \tag{4.12}$$

Consider the function $\varphi(a) = 384\cdot a - (c+\lambda\arctan a)$, with $0<\lambda\le 384$ and $a\in[0,\infty)$. Obviously, $\varphi(0) = -c < 0$ and, as $\lim_{a\to\infty}\varphi(a) = \infty$, we can find $a_0 > 0$ such that $\varphi(a_0) > 0$. This means that

$$c+\lambda \arctan a_0 \le 384 a_0. \tag{4.13}$$

On the other hand, we consider the function $\psi(b) = c + \lambda(\sqrt{2}/2)\arctan(b/16) - 531.61\cdot b$, with $0<\lambda\le 384$ and $b\in[0,\infty)$.

Then, as $\psi(0) = c > 0$ and ψ is a continuous function, we can find b_0 such that

$$\min_{1/4\le t\le 3/4} f\left(t,\frac{b_0}{16}\right) = c+\lambda\frac{\sqrt{2}}{2}\arctan\left(\frac{b_0}{16}\right) \ge b_0\cdot 531.61. \tag{4.14}$$

Therefore, Problem (4.10) can be treated using Theorem 4.1, and we obtain the existence of a symmetric positive solution.

Our main contribution is the uniqueness of the solution.

In what follows, we present the following example which can be treated by Theorem 4.2 and Theorem 4.1 cannot be used.

Example 4.4. Consider the following problem which is a variant of Example 4.3:

$$\begin{gathered}y^{(4)}(t) = c(t) + \lambda\sin(\pi t)\arctan(y(t)), \quad t\in(0,1),\ \lambda>0\\ y(0) = y(1) = y'(0) = y'(1) = 0,\end{gathered} \tag{4.15}$$

where $c(t)$ is a symmetric positive function satisfying $c(1/4) = 0$, for example,

$$c(t) = \begin{cases} 1-4t, & 0\le t\le \dfrac{1}{4},\\ 0, & \dfrac{1}{4}\le t\le \dfrac{3}{4}\\ 4t-3, & \dfrac{3}{4}\le t\le 1.\end{cases} \tag{4.16}$$

In this case, $f(t,y) = c(t) + \lambda\sin(\pi t)\arctan(y(t))$. Taking into account Example 4.3, it is easily proved that $f(t,y)$ satisfies assumptions of Theorem 4.2, and, consequently, Problem (4.15) has a unique symmetric positive solution for $0<\lambda\le 384$.

to Contents, page 1

Now, we prove that $f(t,y)$ does not satisfy assumptions of Theorem 4.1 and, consequently, Problem (4.15) cannot be treated using this theorem. In fact, in this case (notice that $c(1/4) = 0$),

$$\min_{1/4\le t\le 3/4} f\left(t, \frac{b}{16}\right) = f\left(\frac{1}{4}, \frac{b}{16}\right) = \lambda \sin\frac{\pi}{4} \arctan\left(\frac{b}{16}\right) = \lambda\frac{\sqrt{2}}{2}\arctan\left(\frac{b}{16}\right), \tag{4.17}$$

and we cannot find a positive number b_0 such that

$$\lambda\frac{\sqrt{2}}{2}\arctan\left(\frac{b_0}{16}\right) \ge b_0 \cdot 531.61, \quad \text{for } 0 < \lambda \le 384. \tag{4.18}$$

This proves that Problem (4.15) cannot be treated by Theorem 4.1.

Now, we compare our results with the ones obtained in [14]. In [14], the author studies positive solutions of the problem

$$\begin{aligned} &u^{(iv)}(x) = \lambda f(u(x)), \quad x \in (0,1), \\ &u(0) = u(1) = u'(0) = u'(1) = 0, \end{aligned} \tag{4.19}$$

using theory of bifurcation.

His main result works with functions $f(u)$ satisfying

(i) $f(u) > 0$, for $u \le 0$,

(ii) $\lim_{u\to\infty} f(u)/u = \infty$,

(iii) $f'(0) \ge 0$,

(iv) $f''(u) > 0$, for $u > 0$,

and the author proves that there exists a critical λ_0 such that Problem (4.19) has exactly two, exactly one, or no symmetric positive solution depending on whether $0 < \lambda < \lambda_0$, $\lambda = \lambda_0$ or $\lambda > \lambda_0$.

Our Example 3.3 cannot be treated by the results of [14], because, in this case, $f(u) = c + \lambda \arctan u$ and f does not satisfy assumptions (ii) and (iv) above mentioned.

Acknoledgment

This paper is dedicated to Professor Antonio Martin_on on the occasion of his 60th birth-day. This research was partially supported by “Ministerio de Educaci_on y Ciencia”, Project MTM 2007=65706

References

[1] A. Cabada, J. A. Cid, and L. Sánchez, “Positivity and lower and upper solutions for fourth order boundary value problems,” *Nonlinear Analysis*, vol. 67, no. 5, pp. 1599–1612, 2007.

[2] J. Ehme, P. W. Eloe, and J. Henderson, “Upper and lower solution methods for fully nonlinear boundary value problems,” *Journal of Differential Equations*, vol. 180, no. 1, pp. 51–64, 2002.

[3] D. Franco, D. O'Regan, and J. Perán, "Fourth-order problems with nonlinear boundary conditions," *Journal of Computational and Applied Mathematics*, vol. 174, no. 2, pp. 315–327, 2005.
[4] F. Minhós, T. Gyulov, and A. I. Santos, "Existence and location result for a fourth order boundary value problem," *Discrete and Continuous Dynamical System Suplement*, vol. 2005, pp. 662–671, 2005.
[5] Z. Bai, "The method of lower and upper solutions for a bending of an elastic beam equation," *Journal of Mathematical Analysis and Applications*, vol. 248, no. 1, pp. 195–202, 2000.
[6] Y. Guo and Y. Gao, "The method of upper and lower solutions for a Lidstone boundary value problem," *Czechoslovak Mathematical Journal*, vol. 55, no. 3, pp. 639–652, 2005.
[7] P. Habets and L. Sánchez, "A monotone method for fourth order boundary value problems involving a factorizable linear operator," *Portugaliae Mathematica*, vol. 64, no. 3, pp. 255–279, 2007.
[8] J. Caballero, J. Harjani, and K. Sadarangani, "Existence and uniqueness of positive and non-decreasing solutions for a class of singular fractional boundary value problem," *Boundary Value Problems*, vol. 2009, Article ID 421310, 2009.
[9] Z. Bai and H. Wang, "On positive solutions of some nonlinear fourth order beam equations," *Journal of Mathematical Analysis and Applications*, vol. 270, no. 2, pp. 357–368, 2002.
[10] X. L. Liu and W. T. Li, "Existence and multiplicity of solutions for fourth-order boundary value problems with parameters," *Journal of Mathematical Analysis and Applications*, vol. 327, no. 1, pp. 362–375, 2007.
[11] J. R. L. Webb, G. Infante, and D. Franco, "Positive solutions of nonlinear fourth-order boundary-value problems with local and non-local boundary conditions," *Proceedings of the Royal Society of Edinburgh*, vol. 138, no. 2, pp. 427–446, 2008.
[12] C. P. Gupta, "Existence and uniqueness theorems for some fourth order fully quasilinear boundary value problems," *Applicable Analysis*, vol. 36, no. 3-4, pp. 157–169, 1990.
[13] G. E. Hernández and R. Manasevich, "Existence and multiplicity of solutions of a fourth order equation," *Applicable Analysis*, vol. 54, no. 3-4, pp. 237–250, 1994.
[14] P. Korman, "Uniqueness and exact multiplicity of solutions for a class of fourth-order semilinear problems," *Proceedings of the Royal Society of Edinburgh*, vol. 134, no. 1, pp. 179–190, 2004.
[15] R. Ma, "Existence of positive solutions of a fourth-order boundary value problem," *Applied Mathematics and Computation*, vol. 168, no. 2, pp. 1219–1231, 2005.
[16] B. P. Rynne, "Infinitely many solutions of superlinear fourth order boundary value problems," *Topological Methods in Nonlinear Analysis*, vol. 19, no. 2, pp. 303–312, 2002.
[17] A. Amini-Harandi and H. Emami, "A fixed point theorem for contraction type maps in partially ordered metric spaces and application to ordinary differential equations," *Nonlinear Analysis*, vol. 72, no. 5, pp. 2238–2242, 2010.
[18] J. J. Nieto and R. Rodríguez-López, "Contractive mapping theorems in partially ordered sets and applications to ordinary differential equations," *Order*, vol. 22, no. 3, pp. 223–239, 2005.
[19] M. Pei and S. K. Chang, "Monotone iterative technique and symmetric positive solutions for a fourth-order boundary value problem," *Mathematical and Computer Modelling*, vol. 51, no. 9-10, pp. 1260–1267, 2010.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 370982, 14 pages
doi:10.1155/2011/370982

Research Article

Weighted Asymptotically Periodic Solutions of Linear Volterra Difference Equations

Josef Diblík,[1,2] Miroslava Růžičková,[3] Ewa Schmeidel,[4] and Małgorzata Zbąszyniak[4]

[1] *Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Brno University of Technology, 66237 Brno, Czech Republic*
[2] *Department of Mathematics, Faculty of Electrical Engineering and Communication, Brno University of Technology, 61600 Brno, Czech Republic*
[3] *Department of Mathematics, University of Žilina, 01026 Žilina, Slovakia*
[4] *Faculty of Electrical Engineering, Institute of Mathematics, Poznań University of Technology, 60965 Poznań, Poland*

Correspondence should be addressed to Ewa Schmeidel, ewa.schmeidel@put.poznan.pl

Received 16 January 2011; Accepted 17 March 2011

Academic Editor: Elena Braverman

A linear Volterra difference equation of the form $x(n+1) = a(n) + b(n)x(n) + \sum_{i=0}^{n} K(n,i)x(i)$, where $x : \mathbb{N}_0 \to \mathbb{R}$, $a : \mathbb{N}_0 \to \mathbb{R}$, $K : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{R}$ and $b : \mathbb{N}_0 \to \mathbb{R} \setminus \{0\}$ is ω-periodic, is considered. Sufficient conditions for the existence of weighted asymptotically periodic solutions of this equation are obtained. Unlike previous investigations, no restriction on $\prod_{j=0}^{\omega-1} b(j)$ is assumed. The results generalize some of the recent results.

1. Introduction

In the paper, we study a linear Volterra difference equation

$$x(n+1) = a(n) + b(n)x(n) + \sum_{i=0}^{n} K(n,i)x(i), \tag{1.1}$$

where $n \in \mathbb{N}_0 := \{0, 1, 2, \dots\}$, $a : \mathbb{N}_0 \to \mathbb{R}$, $K : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{R}$, and $b : \mathbb{N}_0 \to \mathbb{R} \setminus \{0\}$ is ω-periodic, $\omega \in \mathbb{N} := \{1, 2, \dots\}$. We will also adopt the customary notations

$$\sum_{i=k+s}^{k} \mathcal{O}(i) = 0, \qquad \prod_{i=k+s}^{k} \mathcal{O}(i) = 1, \tag{1.2}$$

to Contents, page 1

where k is an integer, s is a positive integer, and "$\mathcal{O}$" denotes the function considered independently of whether it is defined for the arguments indicated or not.

In [1], the authors considered (1.1) under the assumption

$$\prod_{j=0}^{\omega-1} b(j) = 1, \tag{1.3}$$

and gave sufficient conditions for the existence of asymptotically ω-periodic solutions of (1.1) where the notion for an asymptotically ω-periodic function has been given by the following definition.

Definition 1.1. Let ω be a positive integer. The sequence $y : \mathbb{N}_0 \to \mathbb{R}$ is called ω-periodic if $y(n+\omega) = y(n)$ for all $n \in \mathbb{N}_0$. The sequence y is called asymptotically ω-periodic if there exist two sequences $u, v : \mathbb{N}_0 \to \mathbb{R}$ such that u is ω-periodic, $\lim_{n\to\infty} v(n) = 0$, and

$$y(n) = u(n) + v(n) \tag{1.4}$$

for all $n \in \mathbb{N}_0$.

In this paper, in general, we do not assume that (1.3) holds. Then, we are able to derive sufficient conditions for the existence of a weighted asymptotically ω-periodic solution of (1.1). We give a definition of a weighted asymptotically ω-periodic function.

Definition 1.2. Let ω be a positive integer. The sequence $y : \mathbb{N}_0 \to \mathbb{R}$ is called weighted asymptotically ω-periodic if there exist two sequences $u, v : \mathbb{N}_0 \to \mathbb{R}$ such that u is ω-periodic and $\lim_{n\to\infty} v(n) = 0$, and, moreover, if there exists a sequence $w : \mathbb{N}_0 \to \mathbb{R} \setminus \{0\}$ such that

$$\frac{y(n)}{w(n)} = u(n) + v(n), \tag{1.5}$$

for all $n \in \mathbb{N}_0$.

Apart from this, when we assume

$$\prod_{k=0}^{\omega-1} b(k) = -1, \tag{1.6}$$

then, as a consequence of our main result (Theorem 2.2), the existence of an asymptotically 2ω-periodic solution of (1.1) is obtained.

For the reader's convenience, we note that the background for discrete Volterra equations can be found, for example, in the well-known monograph by Agarwal [2], as well as by Elaydi [3] or Kocić and Ladas [4]. Volterra difference equations were studied by many others, for example, by Appleby et al. [5], by Elaydi and Murakami [6], by Győri and Horváth [7], by Győri and Reynolds [8], and by Song and Baker [9]. For some results on periodic solutions of difference equations, see, for example, [2–4, 10–13] and the related references therein.

to Contents, page 1

2. Weighted Asymptotically Periodic Solutions

In this section, sufficient conditions for the existence of weighted asymptotically ω-periodic solutions of (1.1) will be derived. The following version of Schauder's fixed point theorem given in [14] will serve as a tool used in the proof.

Lemma 2.1. *Let Ω be a Banach space and S its nonempty, closed, and convex subset and let T be a continuous mapping such that $T(S)$ is contained in S and the closure $\overline{T(S)}$ is compact. Then, T has a fixed point in S.*

We set

$$\beta(n) := \prod_{j=0}^{n-1} b(j), \quad n \in \mathbb{N}_0, \tag{2.1}$$

$$\mathcal{B} := \beta(\omega). \tag{2.2}$$

Moreover, we define

$$n^* := n - 1 - \omega \left\lfloor \frac{n-1}{\omega} \right\rfloor, \tag{2.3}$$

where $\lfloor \cdot \rfloor$ is the floor function (the greatest-integer function) and n^* is the "remainder" of dividing $n-1$ by ω. Obviously, $\{\beta(n^*)\}$, $n \in \mathbb{N}$ is an ω-periodic sequence.

Now, we derive sufficient conditions for the existence of a weighted asymptotically ω-periodic solution of (1.1).

Theorem 2.2 (Main result). *Let ω be a positive integer, $b : \mathbb{N}_0 \to \mathbb{R} \setminus \{0\}$ be ω-periodic, $a : \mathbb{N}_0 \to \mathbb{R}$, and $K : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{R}$. Assume that*

$$\begin{gathered} \sum_{i=0}^{\infty} \left| \frac{a(i)}{\beta(i+1)} \right| < \infty, \\ \sum_{j=0}^{\infty} \sum_{i=0}^{j} \left| \frac{K(j,i)\beta(i)}{\beta(j+1)} \right| < 1, \end{gathered} \tag{2.4}$$

and that at least one of the real numbers in the left-hand sides of inequalities (2.4) is positive.

Then, for any nonzero constant c, there exists a weighted asymptotically ω-periodic solution $x : \mathbb{N}_0 \to \mathbb{R}$ of (1.1) with $u, v : \mathbb{N}_0 \to \mathbb{R}$ and $w : \mathbb{N}_0 \to \mathbb{R} \setminus \{0\}$ in representation (1.5) such that

$$w(n) = \mathcal{B}^{\lfloor (n-1)/\omega \rfloor}, \qquad u(n) := c\beta(n^* + 1), \qquad \lim_{n \to \infty} v(n) = 0, \tag{2.5}$$

that is,

$$\frac{x(n)}{\mathcal{B}^{\lfloor (n-1)/\omega \rfloor}} = c\beta(n^* + 1) + v(n), \quad n \in \mathbb{N}_0. \tag{2.6}$$

to Contents, page 1

Proof. We will use a notation

$$M := \sum_{j=0}^{\infty} \sum_{i=0}^{j} \left| \frac{K(j,i)\beta(i)}{\beta(j+1)} \right|, \tag{2.7}$$

whenever this is useful.

Case 1. First assume $c > 0$. We will define an auxiliary sequence of positive numbers $\{\alpha(n)\}$, $n \in \mathbb{N}_0$. We set

$$\alpha(0) := \frac{\sum_{i=0}^{\infty} |a(i)/(\beta(i+1))| + c \sum_{j=0}^{\infty} \sum_{i=0}^{j} |(K(j,i)\beta(i))/(\beta(j+1))|}{1 - \sum_{j=0}^{\infty} \sum_{i=0}^{j} |(K(j,i)\beta(i))/(\beta(j+1))|}, \tag{2.8}$$

where the expression on the right-hand side is well defined due to (2.4). Moreover, we define

$$\alpha(n) := \sum_{i=n}^{\infty} \left| \frac{a(i)}{\beta(i+1)} \right| + (c + \alpha(0)) \sum_{j=n}^{\infty} \sum_{i=0}^{j} \left| \frac{K(j,i)\beta(i)}{\beta(j+1)} \right|, \tag{2.9}$$

for $n \geq 1$. It is easy to see that

$$\lim_{n \to \infty} \alpha(n) = 0. \tag{2.10}$$

We show, moreover, that

$$\alpha(n) \leq \alpha(0), \tag{2.11}$$

for any $n \in \mathbb{N}$. Let us first remark that

$$\alpha(0) = \sum_{i=0}^{\infty} \left| \frac{a(i)}{\beta(i+1)} \right| + (c + \alpha(0)) \sum_{j=0}^{\infty} \sum_{i=0}^{j} \left| \frac{K(j,i)\beta(i)}{\beta(j+1)} \right|. \tag{2.12}$$

Then, due to the convergence of both series (see (2.4)), the inequality

$$\begin{aligned} \alpha(0) &= \sum_{i=0}^{\infty} \left| \frac{a(i)}{\beta(i+1)} \right| + (c + \alpha(0)) \sum_{j=0}^{\infty} \sum_{i=0}^{j} \left| \frac{K(j,i)\beta(i)}{\beta(j+1)} \right| \\ &\geq \sum_{i=n}^{\infty} \left| \frac{a(i)}{\beta(i+1)} \right| + (c + \alpha(0)) \sum_{j=n}^{\infty} \sum_{i=0}^{j} \left| \frac{K(j,i)\beta(i)}{\beta(j+1)} \right| = \alpha(n) \end{aligned} \tag{2.13}$$

obviously holds for every $n \in \mathbb{N}$ and (2.11) is proved.

to Contents, page 1

Let B be the Banach space of all real bounded sequences $z : \mathbb{N}_0 \to \mathbb{R}$ equipped with the usual supremum norm $\|z\| = \sup_{n\in\mathbb{N}_0}|z(n)|$ for $z \in B$. We define a subset $S \subset B$ as

$$S := \{z \in B : c - \alpha(0) \le z(n) \le c + \alpha(0),\ n \in \mathbb{N}_0\}. \tag{2.14}$$

It is not difficult to prove that S is a nonempty, bounded, convex, and closed subset of B.

Let us define a mapping $T : S \to B$ as follows:

$$(Tz)(n) = c - \sum_{i=n}^{\infty} \frac{a(i)}{\beta(i+1)} - \sum_{j=n}^{\infty}\sum_{i=0}^{j} \frac{K(j,i)\beta(i)}{\beta(j+1)} z(i), \tag{2.15}$$

for any $n \in \mathbb{N}_0$.

We will prove that the mapping T has a fixed point in S.

We first show that $T(S) \subset S$. Indeed, if $z \in S$, then $|z(n) - c| \le \alpha(0)$ for $n \in \mathbb{N}_0$ and, by (2.11) and (2.15), we have

$$|(Tz)(n) - c| \le \sum_{i=n}^{\infty} \left|\frac{a(i)}{\beta(i+1)}\right| + (c + \alpha(0)) \sum_{j=n}^{\infty}\sum_{i=0}^{j} \left|\frac{K(j,i)\beta(i)}{\beta(j+1)}\right| = \alpha(n) \le \alpha(0). \tag{2.16}$$

Next, we prove that T is continuous. Let $z^{(p)}$ be a sequence in S such that $z^{(p)} \to z$ as $p \to \infty$. Because S is closed, $z \in S$. Now, utilizing (2.15), we get

$$\begin{aligned} \left|\left(Tz^{(p)}\right)(n) - (Tz)(n)\right| &= \left|\sum_{j=n}^{\infty}\sum_{i=0}^{j} \frac{K(j,i)\beta(i)}{\beta(j+1)} \left(z^{(p)}(i) - z(i)\right)\right| \\ &\le M \sup_{i\ge 0}\left|z^{(p)}(i) - z(i)\right| = M\left\|z^{(p)} - z\right\|, \quad n \in \mathbb{N}_0. \end{aligned} \tag{2.17}$$

Therefore,

$$\begin{gathered} \left\|Tz^{(p)} - Tz\right\| \le M\left\|z^{(p)} - z\right\|, \\ \lim_{p\to\infty}\left\|Tz^{(p)} - Tz\right\| = 0. \end{gathered} \tag{2.18}$$

This means that T is continuous.

Now, we show that $\overline{T(S)}$ is compact. As is generally known, it is enough to verify that every ε-open covering of $\overline{T(S)}$ contains a finite ε-subcover of $\overline{T(S)}$, that is, finitely many of these open sets already cover $\overline{T(S)}$ ([15], page 756 (12)). Thus, to prove that $\overline{T(S)}$ is compact, we take an arbitrary $\varepsilon > 0$ and assume that an open ε-cover $\mathcal{C}_\varepsilon$ of $\overline{T(S)}$ is given. Then, from (2.10), we conclude that there exists an $n_\varepsilon \in \mathbb{N}$ such that $\alpha(n) < \varepsilon/4$ for $n \ge n_\varepsilon$.

to Contents, page 1

Suppose that $x_T^1 \in \overline{T(S)}$ is one of the elements generating the ε-cover $\mathcal{C}_\varepsilon$ of $\overline{T(S)}$. Then (as follows from (2.16)), for an arbitrary $x_T \in \overline{T(\mathrm{S})}$,

$$\left| x_T^1(n) - x_T(n) \right| < \varepsilon \tag{2.19}$$

if $n \geq n_\varepsilon$. In other words, the ε-neighborhood of $x_T^1 - c^*$:

$$\left\| x_T^1 - c^* \right\| < \varepsilon, \tag{2.20}$$

where $c^* = \{c, c, \ldots\} \in S$ covers the set $\overline{T(S)}$ on an infinite interval $n \geq n_\varepsilon$. It remains to cover the rest of $\overline{T(S)}$ on a finite interval for $n \in \{0, 1 \ldots, n_\varepsilon - 1\}$ by a finite number of ε-neighborhoods of elements generating ε-cover $\mathcal{C}_\varepsilon$. Supposing that x_T^1 itself is not able to generate such cover, we fix $n \in \{0, 1, \ldots, n_\varepsilon - 1\}$ and split the interval

$$[c - \alpha(n), c + \alpha(n)] \tag{2.21}$$

into a finite number $h(\varepsilon, n)$ of closed subintervals

$$I_1(n), I_2(n), \ldots, I_{h(\varepsilon,n)}(n) \tag{2.22}$$

each with a length not greater then $\varepsilon/2$ such that

$$\begin{gathered} \bigcup_{i=1}^{h(\varepsilon,n)} I_i(n) = [c - \alpha(n), c + \alpha(n)], \\ \operatorname{int} I_i(n) \cap \operatorname{int} I_j(n) = \emptyset, \quad i, j = 1, 2, \ldots, h(\varepsilon, n), \ i \neq j. \end{gathered} \tag{2.23}$$

Finally, the set

$$\bigcup_{n=0}^{n_\varepsilon - 1} [c - \alpha(n), c + \alpha(n)] \tag{2.24}$$

equals

$$\bigcup_{n=0}^{n_\varepsilon - 1} \bigcup_{i=1}^{h(\varepsilon,n)} I_i(n) \tag{2.25}$$

and can be divided into a finite number

$$M_\varepsilon := \sum_{n=0}^{n_\varepsilon - 1} h(\varepsilon, n) \tag{2.26}$$

to Contents, page 1

of different subintervals (2.22). This means that, at most, M_ε of elements generating the cover $\mathcal{C}_\varepsilon$ are sufficient to generate a finite ε-subcover of $\overline{T(S)}$ for $n \in \{0, 1, \ldots, n_\varepsilon - 1\}$. We remark that each of such elements simultaneously plays the same role as $x_T^1(n)$ for $n \geq n_\varepsilon$. Since $\varepsilon > 0$ can be chosen as arbitrarily small, $\overline{T(S)}$ is compact.

By Schauder's fixed point theorem, there exists a $z \in S$ such that $z(n) = (Tz)(n)$ for $n \in \mathbb{N}_0$. Thus,

$$z(n) = c - \sum_{i=n}^{\infty} \frac{a(i)}{\beta(i+1)} - \sum_{j=n}^{\infty} \sum_{i=0}^{j} \frac{\beta(i)}{\beta(j+1)} K(j,i) z(i), \tag{2.27}$$

for any $n \in \mathbb{N}_0$.

Due to (2.10) and (2.16), for fixed point $z \in S$ of T, we have

$$\lim_{n\to\infty} |z(n) - c| = \lim_{n\to\infty} |(Tz)(n) - c| \leq \lim_{n\to\infty} \alpha(n) = 0, \tag{2.28}$$

or, equivalently,

$$\lim_{n\to\infty} z(n) = c. \tag{2.29}$$

Finally, we will show that there exists a connection between the fixed point $z \in S$ and the existence of a solution of (1.1) which divided by $\mathcal{B}^{\lfloor (n-1)/\omega \rfloor}$ provides an asymptotically ω-periodic sequence. Considering (2.27) for $z(n+1)$ and $z(n)$, we get

$$\Delta z(n) = \frac{a(n)}{\beta(n+1)} + \sum_{i=0}^{n} \frac{\beta(i)}{\beta(n+1)} K(n,i) z(i), \tag{2.30}$$

where $n \in \mathbb{N}_0$. Hence, we have

$$z(n+1) - z(n) = \frac{a(n)}{\beta(n+1)} + \frac{1}{\beta(n+1)} \sum_{i=0}^{n} \beta(i) K(n,i) z(i), \quad n \in \mathbb{N}_0. \tag{2.31}$$

Putting

$$z(n) = \frac{x(n)}{\beta(n)}, \quad n \in \mathbb{N}_0 \tag{2.32}$$

in (2.31), we get (1.1) since

$$\frac{x(n+1)}{\beta(n+1)} - \frac{x(n)}{\beta(n)} = \frac{a(n)}{\beta(n+1)} + \frac{1}{\beta(n+1)} \sum_{i=0}^{n} K(n,i) x(i), \quad n \in \mathbb{N}_0 \tag{2.33}$$

to Contents, page 1

yields

$$x(n+1) = a(n) + b(n)x(n) + \sum_{i=0}^{n} K(n,i)x(i), \quad n \in \mathbb{N}_0. \tag{2.34}$$

Consequently, x defined by (2.32) is a solution of (1.1). From (2.29) and (2.32), we obtain

$$\frac{x(n)}{\beta(n)} = z(n) = c + o(1), \tag{2.35}$$

for $n \to \infty$ (where $o(1)$ is the Landau order symbol). Hence,

$$x(n) = \beta(n)(c + o(1)), \quad n \longrightarrow \infty. \tag{2.36}$$

It is easy to show that the function β defined by (2.1) can be expressed in the form

$$\beta(n) = \prod_{j=0}^{n-1} b(j) = \mathcal{B}^{\lfloor (n-1)/\omega \rfloor} \cdot \beta(n^* + 1), \tag{2.37}$$

for $n \in \mathbb{N}_0$. Then, as follows from (2.36),

$$x(n) = \mathcal{B}^{\lfloor (n-1)/\omega \rfloor} \cdot \beta(n^* + 1)(c + o(1)), \quad n \longrightarrow \infty, \tag{2.38}$$

or

$$\frac{x(n)}{\mathcal{B}^{\lfloor (n-1)/\omega \rfloor}} = c\beta(n^* + 1) + \beta(n^* + 1)o(1), \quad n \longrightarrow \infty. \tag{2.39}$$

The proof is completed since the sequence $\{\beta(n^* + 1)\}$ is ω-periodic, hence bounded and, due to the properties of Landau order symbols, we have

$$\beta(n^* + 1)o(1) = o(1), \quad n \longrightarrow \infty, \tag{2.40}$$

and it is easy to see that the choice

$$u(n) := c\beta(n^* + 1), \qquad w(n) := \mathcal{B}^{\lfloor (n-1)/\omega \rfloor}, \quad n \in \mathbb{N}_0, \tag{2.41}$$

and an appropriate function $v : \mathbb{N}_0 \to \mathbb{R}$ such that

$$\lim_{n \to \infty} v(n) = 0 \tag{2.42}$$

finishes this part of the proof. Although for $n = 0$, there is no correspondence between formula (2.36) and the definitions of functions u and w, we assume that function v makes up for this.

to Contents, page 1

Case 2. If $c < 0$, we can proceed as follows. It is easy to see that arbitrary solution $y = y(n)$ of the equation

$$y(n+1) = -a(n) + b(n)y(n) + \sum_{i=0}^{n} K(n,i)y(i) \tag{2.43}$$

defines a solution $x = x(n)$ of (1.1) since a substitution $y(n) = -x(n)$ in (2.43) turns (2.43) into (1.1). If the assumptions of Theorem 2.2 hold for (1.1), then, obviously, Theorem 2.2 holds for (2.43) as well. So, for an arbitrary $c > 0$, (2.43) has a solution that can be represented by formula (2.6), that is,

$$\frac{y(n)}{\mathcal{B}^{\lfloor (n-1)/\omega \rfloor}} = c\beta(n^* + 1) + v(n), \quad n \in \mathbb{N}_0. \tag{2.44}$$

Or, in other words, (1.1) has a solution that can be represented by formula (2.44) as

$$\frac{x(n)}{\mathcal{B}^{\lfloor (n-1)/\omega \rfloor}} = c_0\beta(n^* + 1) + v^*(n), \quad n \in \mathbb{N}_0, \tag{2.45}$$

with $c_0 = -c$ and $v^*(n) = -v(n)$. In (2.45), $c_0 < 0$ and the function $v^*(n)$ has the same properties as the function $v(n)$. Therefore, formula (2.6) is valid for an arbitrary negative c as well. □

Now, we give an example which illustrates the case where there exists a solution to equation of the type (1.1) which is weighted asymptotically periodic, but is not asymptotically periodic.

Example 2.3. We consider (1.1) with

$$\begin{gathered} a(n) = (-1)^{n+1}\left(1 - \frac{1}{3^{n+1}}\right), \\ b(n) = 3(-1)^n, \\ K(n,i) = (-1)^{n+(i(i-1))/2}\frac{1}{3^{2i}}, \end{gathered} \tag{2.46}$$

that is, the equation

$$x(n+1) = (-1)^{n+1}\left(1 - \frac{1}{3^{n+1}}\right) + 3(-1)^n x(n) + \sum_{i=0}^{n}(-1)^{n+(i(i-1))/2}\frac{1}{3^{2i}}x(i). \tag{2.47}$$

to Contents, page 1

The sequence $b(n)$ is 2-periodic and

$$\beta(n) = \prod_{j=0}^{n-1} b(j) = (-1)^{n(n-1)/2} 3^n ,$$

$$\mathcal{B} = \beta(\omega) = \beta(2) = -9,$$

$$\beta(n^* + 1) = -3 + 6(-1)^{n+1},$$

$$\frac{a(n)}{\beta(n+1)} = (-1)^{(-n^2+n+2)/2}\left(\frac{1}{3^{n+1}} - \frac{1}{3^{2(n+1)}}\right),$$

$$\sum_{i=0}^{\infty}\left|\frac{a(i)}{\beta(i+1)}\right| < \infty, \tag{2.48}$$

$$\begin{aligned}\sum_{j=0}^{\infty}\sum_{i=0}^{j}\left|\frac{K(j,i)\beta(i)}{\beta(j+1)}\right| &< \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\left|\frac{K(j,i)\beta(i)}{\beta(j+1)}\right| = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\frac{1}{3^{i+j+1}} \\ &= \frac{1}{3}\left(\sum_{j=0}^{\infty}\frac{1}{3^j}\right)\left(\sum_{i=0}^{\infty}\frac{1}{3^i}\right) = \frac{1}{3}\cdot\frac{1}{1-1/3}\cdot\frac{1}{1-1/3} \\ &= \frac{1}{3}\cdot\frac{3}{2}\cdot\frac{3}{2} = \frac{3}{4} < 1.\end{aligned}$$

By virtue of Theorem 2.2, for any nonzero constant c, there exists a solution $x : \mathbb{N}_0 \to \mathbb{R}$ of (1.1) which is weighed asymptotically 2-periodic. Let, for example, $c = 2/3$. Then,

$$\begin{gathered} w(n) = (-9)^{\lfloor (n-1)/2 \rfloor}, \\ u(n) = c\beta(n^* + 1) = \frac{2}{3}\left(-3 + 6(-1)^{n+1}\right) = -2 + 4(-1)^{n+1}, \end{gathered} \tag{2.49}$$

and the sequence $x(n)$ given by

$$\frac{x(n)}{(-9)^{\lfloor (n-1)/2 \rfloor}} = -2 + 4(-1)^{n+1} + v(n), \quad n \in \mathbb{N}_0, \tag{2.50}$$

or, equivalently,

$$x(n) = (-9)^{\lfloor (n-1)/2 \rfloor}\left(-2 + 4(-1)^{n+1}\right) + v(n), \quad n \in \mathbb{N}_0 \tag{2.51}$$

is such a solution. We remark that such solution is not asymptotically 2-periodic in the meaning of Definition 1.1.

to Contents, page 1

It is easy to verify that the sequence $x^*(n)$ obtained from (2.51) if $v(n) = 0$, $n \in \mathbb{N}_0$, that is,

$$x^*(n) = (-9)^{\lfloor (n-1)/2 \rfloor}\left(-2 + 4(-1)^{n+1}\right) = \frac{2}{3} \cdot (-1)^{n(n-1)/2} \cdot 3^n, \quad n \in \mathbb{N}_0 \tag{2.52}$$

is a true solution of (2.47).

3. Concluding Remarks and Open Problems

It is easy to prove the following corollary.

Corollary 3.1. *Let Theorem 2.2 be valid. If, moreover, $|\mathcal{B}| < 1$, then every solution $x = x(n)$ of* (1.1) *described by formula* (2.6) *satisfies*

$$\lim_{n\to\infty} x(n) = 0. \tag{3.1}$$

If $|\mathcal{B}| > 1$, then, for every solution $x = x(n)$ of (1.1) *described by formula* (2.6), *one has*

$$\liminf_{n\to\infty} x(n) = -\infty \tag{3.2}$$

or/and

$$\limsup_{n\to\infty} x(n) = \infty. \tag{3.3}$$

Finally, if $\mathcal{B} > 1$, then, for every solution $x = x(n)$ of (1.1) *described by formula* (2.6), *one has*

$$\lim_{n\to\infty} x(n) = \infty, \tag{3.4}$$

and if $\mathcal{B} < -1$, then, for every solution $x = x(n)$ of (1.1) *described by formula* (2.6), *one has*

$$\lim_{n\to\infty} x(n) = -\infty. \tag{3.5}$$

Now, let us discuss the case when (1.6) holds, that is, when

$$\mathcal{B} = \prod_{j=0}^{\omega-1} b(j) = -1. \tag{3.6}$$

Corollary 3.2. *Let Theorem 2.2 be valid. Assume that $\mathcal{B} = -1$. Then, for any nonzero constant c, there exists an asymptotically 2ω-periodic solution $x = x(n)$, $n \in \mathbb{N}_0$ of* (1.1) *such that*

$$x(n) = (-1)^{\lfloor (n-1)/\omega \rfloor} u(n) + z(n), \quad n \in \mathbb{N}_0, \tag{3.7}$$

to Contents, page 1

with

$$u(n) := c\beta(n^* + 1), \qquad \lim_{n\to\infty} z(n) = 0. \tag{3.8}$$

Proof. Putting $\mathcal{B} = -1$ in Theorem 2.2, we get

$$x(n) = (-1)^{\lfloor (n-1)/\omega \rfloor} u(n) + (-1)^{\lfloor (n-1)/\omega \rfloor} v(n), \tag{3.9}$$

with

$$u(n) := c\beta(n^* + 1), \qquad \lim_{n\to\infty} v(n) = 0. \tag{3.10}$$

Due to the definition of n^*, we see that the sequence

$$\left\{\beta(n^* + 1)\right\} = \{\beta(\omega), \beta(1), \beta(2), \dots, \beta(\omega), \beta(1), \beta(2), \dots, \beta(\omega), \dots\}, \tag{3.11}$$

is an ω-periodic sequence. Since

$$\left\{ \left\lfloor \frac{n-1}{\omega} \right\rfloor \right\} = \left\{ -1, \underbrace{0, \dots, 0}_{\omega}, \underbrace{1, \dots, 1}_{\omega}, 2, \dots \right\}, \tag{3.12}$$

for $n \in \mathbb{N}_0$, we have

$$\left\{ (-1)^{\lfloor (n-1)/\omega \rfloor} \right\} = \left\{ -1, \underbrace{1, \dots, 1}_{\omega}, \underbrace{-1, \dots, -1}_{\omega}, 1, \dots \right\}. \tag{3.13}$$

Therefore, the sequence

$$\left\{ (-1)^{\lfloor (n-1)/\omega \rfloor} u(n) \right\} = c\{-\beta(\omega), \beta(1), \beta(2), \dots, \beta(\omega), -\beta(1), -\beta(2), \dots, -\beta(\omega), \dots\} \tag{3.14}$$

is a 2ω-periodic sequence. Set

$$z(n) = (-1)^{\lfloor (n-1)/\omega \rfloor} v(n). \tag{3.15}$$

Then,

$$\lim_{n\to\infty} z(n) = 0. \tag{3.16}$$

The proof is completed. □

to Contents, page 1

Remark 3.3. From the proof, we see that Theorem 2.2 remains valid even in the case of $c = 0$. Then, there exists an "asymptotically weighted ω-periodic solution" $x = x(n)$ of (1.1) as well. The formula (2.6) reduces to

$$x(n) = \mathcal{B}^{\lfloor (n-1)/\omega \rfloor} v(n) = o(1), \quad n \in \mathbb{N}_0, \tag{3.17}$$

since $u(n) = 0$. In the light of Definition 1.2, we can treat this case as follows. We set (as a singular case) $u \equiv 0$ with an arbitrary (possibly other than "ω") period and with $v = o(1)$, $n \to \infty$.

Remark 3.4. The assumptions of Theorem 2.2 [1] are substantially different from those of the present Theorem 2.2. However, it is easy to see that Theorem 2.2 [1] is a particular case of the present Theorem 2.2 if (1.3) holds, that is, if $\mathcal{B} = 1$. Therefore, our results can be viewed as a generalization of some results in [1].

In connection with the above investigations, some open problems arise.

Open Problem 1. The results of [1] are extended to systems of linear Volterra discrete equations in [16, 17]. It is an open question if the results presented can be extended to systems of linear Volterra discrete equations.

Open Problem 2. Unlike the result of Theorem 2.2 [1] where a parameter c can be arbitrary, the assumptions of the results in [16, 17] are more restrictive since the related parameters should satisfy certain inequalities as well. Different results on the existence of asymptotically periodic solutions were recently proved in [8]. Using an example, it is shown that the results in [8] can be less restrictive. Therefore, an additional open problem arises if the results in [16, 17] can be improved in such a way that the related parameters can be arbitrary and if the expected extension of the results suggested in Open Problem 1 can be given in such a way that the related parameters can be arbitrary as well.

Acknowledgments

The first author has been supported by the Grant P201/10/1032 of the Czech Grant Agency (Prague), by the Council of Czech Government MSM 00216 30519, and by the project FEKT/FSI-S-11-1-1159. The second author has been supported by the Grant VEGA 1/0090/09 of the Grant Agency of Slovak Republic and by the Grant APVV-0700-07 of the Slovak Research and Development Agency.

References

[1] J. Diblík, M. Růžičková, and E. Schmeidel, "Asymptotically periodic solutions of Volterra difference equations," *Tatra Mountains Mathematical Publications*, vol. 43, pp. 43–61, 2009.

[2] R. P. Agarwal, *Difference Equations and Inequalities*, vol. 228 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 2nd edition, 2000.

[3] S. N. Elaydi, *An Introduction to Difference Equations*, Undergraduate Texts in Mathematics, Springer, New York, NY, USA, 3rd edition, 2005.

[4] V. L. Kocić and G. Ladas, *Global Behavior of Nonlinear Difference Equations of Higher Order with Applications*, vol. 256 of *Mathematics and Its Applications*, Kluwer Academic, Dordrecht, The Netherlands, 1993.

to Contents, page 1

[5] J. A. D. Appleby, I. Győri, and D. W. Reynolds, "On exact convergence rates for solutions of linear systems of Volterra difference equations," *Journal of Difference Equations and Applications*, vol. 12, no. 12, pp. 1257–1275, 2006.
[6] S. Elaydi and S. Murakami, "Uniform asymptotic stability in linear Volterra difference equations," *Journal of Difference Equations and Applications*, vol. 3, no. 3-4, pp. 203–218, 1998.
[7] I. Győri and L. Horváth, "Asymptotic representation of the solutions of linear Volterra difference equations," *Advances in Difference Equations*, vol. 2008, Article ID 932831, 22 pages, 2008.
[8] I. Győri and D. W. Reynolds, "On asymptotically periodic solutions of linear discrete Volterra equations," *Fasciculi Mathematici*, no. 44, pp. 53–67, 2010.
[9] Y. Song and C. T. H. Baker, "Admissibility for discrete Volterra equations," *Journal of Difference Equations and Applications*, vol. 12, no. 5, pp. 433–457, 2006.
[10] L. Berg and S. Stević, "Periodicity of some classes of holomorphic difference equations," *Journal of Difference Equations and Applications*, vol. 12, no. 8, pp. 827–835, 2006.
[11] S. Stević, "On global periodicity of a class of difference equations," *Discrete Dynamics in Nature and Society*, vol. 2007, Article ID 23503, 10 pages, 2007.
[12] S. Stević, "Periodicity of max difference equations," *Utilitas Mathematica*, vol. 83, pp. 69–71, 2010.
[13] S. Stević and K. S. Berenhaut, "The behavior of positive solutions of a nonlinear second-order difference equation $x_n = f(x_{n-2})/g(x_{n-1})$," *Abstract and Applied Analysis*, vol. 2008, Article ID 653243, 8 pages, 2008.
[14] J. Musielak, *Wstep Do Analizy Funkcjonalnej*, PWN, Warszawa, Poland, 1976.
[15] E. Zeidler, *Nonlinear Functional Analysis and Its Applications. I*, Springer, New York, NY, USA, 1986.
[16] J. Diblík, E. Schmeidel, and M. Růžičková, "Asymptotically periodic solutions of Volterra system of difference equations," *Computers & Mathematics with Applications*, vol. 59, no. 8, pp. 2854–2867, 2010.
[17] J. Diblík, E. Schmeidel, and M. Růžičková, "Existence of asymptotically periodic solutions of system of Volterra difference equations," *Journal of Difference Equations and Applications*, vol. 15, no. 11-12, pp. 1165–1177, 2009.

to Contents, page 1

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2011, Article ID 738520, 41 pages
doi:10.1155/2011/738520

Research Article

Weyl-Titchmarsh Theory for Time Scale Symplectic Systems on Half Line

Roman Šimon Hilscher and Petr Zemánek

Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Kotlářská 2, 61137 Brno, Czech Republic

Correspondence should be addressed to Roman Šimon Hilscher, hilscher@math.muni.cz

Received 8 October 2010; Accepted 3 January 2011

Academic Editor: Miroslava Růžičková

We develop the Weyl-Titchmarsh theory for time scale symplectic systems. We introduce the $M(\lambda)$-function, study its properties, construct the corresponding Weyl disk and Weyl circle, and establish their geometric structure including the formulas for their center and matrix radii. Similar properties are then derived for the limiting Weyl disk. We discuss the notions of the system being in the limit point or limit circle case and prove several characterizations of the system in the limit point case and one condition for the limit circle case. We also define the Green function for the associated nonhomogeneous system and use its properties for deriving further results for the original system in the limit point or limit circle case. Our work directly generalizes the corresponding discrete time theory obtained recently by S. Clark and P. Zemánek (2010). It also unifies the results in many other papers on the Weyl-Titchmarsh theory for linear Hamiltonian differential, difference, and dynamic systems when the spectral parameter appears in the second equation. Some of our results are new even in the case of the second-order Sturm-Liouville equations on time scales.

1. Introduction

In this paper we develop systematically the Weyl-Titchmarsh theory for time scale symplectic systems. Such systems unify and extend the classical linear Hamiltonian differential systems and discrete symplectic and Hamiltonian systems, including the Sturm-Liouville differential and difference equations of arbitrary even order. As the research in the Weyl-Titchmarsh theory has been very active in the last years, we contribute to this development by presenting a theory which directly generalizes and unifies the results in several recent papers, such as [1–4] and partly in [5–14].

to Contents, page 1

Historically, the theory nowadays called by Weyl and Titchmarsh started in [15] by the investigation of the second-order linear differential equation

$$\left(r(t)z'(t)\right) + q(t)z(t) = \lambda z(t), \quad t \in [0, \infty), \tag{1.1}$$

where $r, q : [0, \infty) \to \mathbb{R}$ are continuous, $r(t) > 0$, and $\lambda \in \mathbb{C}$, is a spectral parameter. By using a geometrical approach it was showed that (1.1) can be divided into two classes called the limit circle and limit point meaning that either all solutions of (1.1) are square integrable for all $\lambda \in \mathbb{C} \setminus \mathbb{R}$ or there is a unique (up to a multiplicative constant) square-integrable solution of (1.1) on $[0, \infty)$. Analytic methods for the investigation of (1.1) have been introduced in a series of papers starting with [16]; see also [17]. We refer to [18–20] for an overview of the original contributions to the Weyl-Titchmarsh theory for (1.1); see also [21]. Extensions of the Weyl-Titchmarsh theory to more general equations, namely, to the linear Hamiltonian differential systems

$$z'(t) = [\lambda A(t) + B(t)]z(t), \quad t \in [0, \infty), \tag{1.2}$$

was initiated in [22] and developed further in [6, 8, 10, 11, 23–38].

According to [19], the first paper dealing with the parallel discrete time Weyl theory for second-order difference equations appears to be the work mentioned in [39]. Since then a long time elapsed until the theory of difference equations attracted more attention. The Weyl-Titchmarsh theory for the second-order Sturm-Liouville difference equations was developed in [22, 40, 41]; see also the references in [19]. For higher-order Sturm-Liouville difference equations and linear Hamiltonian difference systems, such as

$$\Delta x_k = A_k x_{k+1} + \left(B_k + \lambda W_k^{[2]}\right) u_k, \quad \Delta u_k = \left(C_k - \lambda W_k^{[1]}\right) x_{k+1} - A_k^* u_k, \quad k \in [0, \infty)_{\mathbb{Z}}, \tag{1.3}$$

where A_k, B_k, C_k, $W_k^{[1]}$, $W_k^{[2]}$ are complex $n \times n$ matrices such that B_k and C_k are Hermitian and $W_k^{[1]}$ and $W_k^{[2]}$ are Hermitian and nonnegative definite, the Weyl-Titchmarsh theory was studied in [9, 14, 42]. Recently, the results for linear Hamiltonian difference systems were generalized in [1, 2] to discrete symplectic systems

$$x_{k+1} = \mathcal{A}_k x_k + \mathcal{B}_k u_k, \quad u_{k+1} = \mathcal{C}_k x_k + \mathfrak{D}_k u_k + \lambda \mathcal{W}_k x_{k+1}, \quad k \in [0, \infty)_{\mathbb{Z}}, \tag{1.4}$$

where $\mathcal{A}_k$, $\mathcal{B}_k$, $\mathcal{C}_k$, $\mathfrak{D}_k$, $\mathcal{W}_k$ are complex $n \times n$ matrices such that $\mathcal{W}_k$ is Hermitian and nonnegative definite and the $2n \times 2n$ transition matrix in (1.4) is symplectic, that is,

$$\mathcal{S}_k := \begin{pmatrix} \mathcal{A}_k & \mathcal{B}_k \\ \mathcal{C}_k & \mathfrak{D}_k \end{pmatrix}, \quad \mathcal{S}_k^* \mathcal{J} \mathcal{S}_k = \mathcal{J}, \quad \mathcal{J} := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}. \tag{1.5}$$

In the unifying theory for differential and difference equations—the theory of time scales—the classification of second-order Sturm-Liouville dynamic equations

$$y^{\Delta\Delta}(t) + q(t)y^{\sigma}(t) = \lambda y^{\sigma}(t), \quad t \in [a, \infty)_{\mathbb{T}}, \tag{1.6}$$

to Contents, page 1

to be of the limit point or limit circle type is given in [4, 43]. These two papers seem to be the only ones on time scales which are devoted to the Weyl-Titchmarsh theory for the second order dynamic equations. Another way of generalizing the Weyl-Titchmarsh theory for continuous and discrete Hamiltonian systems was presented in [3, 5]. In these references the authors consider the linear Hamiltonian system

$$\begin{aligned} x^{\Delta}(t) &= A(t)x^{\sigma}(t) + [B(t) + \lambda W_2(t)]u(t), \\ u^{\Delta}(t) &= [C(t) - \lambda W_1(t)]x^{\sigma}(t) - A^*(t)u(t), \quad t \in [a,\infty)_{\mathbb{T}}, \end{aligned} \tag{1.7}$$

on the so-called Sturmian or general time scales, respectively. Here $f^{\Delta}(t)$ is the time scale Δ-derivative and $f^{\sigma}(t) := f(\sigma(t))$, where $\sigma(t)$ is the forward jump at t; see the time scale notation in Section 2.

In the present paper we develop the Weyl-Titchmarsh theory for more general linear dynamic systems, namely, the time scale symplectic systems

$$\begin{aligned} x^{\Delta}(t) &= \mathcal{A}(t)x(t) + \mathcal{B}(t)u(t), \\ u^{\Delta}(t) &= \mathcal{C}(t)x(t) + \mathfrak{D}(t)u(t) - \lambda \mathcal{W}(t)x^{\sigma}(t), \quad t \in [a,\infty)_{\mathbb{T}}, \end{aligned} \tag{$\mathcal{S}_\lambda$}$$

where $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathfrak{D}$, $\mathcal{W}$ are complex $n \times n$ matrix functions on $[a,\infty)_{\mathbb{T}}$, $\mathcal{W}(t)$ is Hermitian and nonnegative definite, $\lambda \in \mathbb{C}$, and the $2n \times 2n$ coefficient matrix in system $(\mathcal{S}_\lambda)$ satisfies

$$\mathcal{S}(t) := \begin{pmatrix} \mathcal{A}(t) & \mathcal{B}(t) \\ \mathcal{C}(t) & \mathfrak{D}(t) \end{pmatrix}, \quad \mathcal{S}^*(t)\mathcal{J} + \mathcal{J}\mathcal{S}(t) + \mu(t)\mathcal{S}^*(t)\mathcal{J}\mathcal{S}(t) = 0, \quad t \in [a,\infty)_{\mathbb{T}}, \tag{1.8}$$

where $\mu(t) := \sigma(t) - t$ is the graininess of the time scale. The spectral parameter λ is only in the second equation of system $(\mathcal{S}_\lambda)$. This system was introduced in [44], and it naturally unifies the previously mentioned continuous, discrete, and time scale linear Hamiltonian systems (having the spectral parameter in the second equation only) and discrete symplectic systems into one framework. Our main results are the properties of the $M(\lambda)$ function, the geometric description of the Weyl disks, and characterizations of the limit point and limit circle cases for the time scale symplectic system $(\mathcal{S}_\lambda)$. In addition, we give a formula for the $L^2_{\mathcal{W}}$ solutions of a nonhomogeneous time scale symplectic system in terms of its Green function. These results generalize and unify in particular all the results in [1–4] and some results from [5–14]. The theory of time scale symplectic systems or Hamiltonian systems is a topic with active research in recent years; see, for example, [44–51]. This paper can be regarded not only as a completion of these papers by establishing the Weyl-Titchmarsh theory for time scale symplectic systems but also as a comparison of the corresponding continuous and discrete time results. The references to particular statements in the literature are displayed throughout the text. Many results of this paper are new even for (1.6), being a special case of system $(\mathcal{S}_\lambda)$. An overview of these new results for (1.6) will be presented in our subsequent work.

This paper is organized as follows. In the next section we recall some basic notions from the theory of time scales and linear algebra. In Section 3 we present fundamental properties of time scale symplectic systems with complex coefficients, including the important Lagrange identity (Theorem 3.5) and other formulas involving their solutions.

to Contents, page 1

In Section 4 we define the time scale $M(\lambda)$-function for system $(\mathcal{S}_\lambda)$ and establish its basic properties in the case of the regular spectral problem. In Section 5 we introduce the Weyl disks and circles for system $(\mathcal{S}_\lambda)$ and describe their geometric structure in terms of contractive matrices in $\mathbb{C}^{n\times n}$. The properties of the limiting Weyl disk and Weyl circle are then studied in Section 6, where we also prove that system $(\mathcal{S}_\lambda)$ has at least n linearly independent solutions in the space $L^2_{\mathcal{W}}$ (see Theorem 6.7). In Section 7 we define the system $(\mathcal{S}_\lambda)$ to be in the limit point and limit circle case and prove several characterizations of these properties. In the final section we consider the system $(\mathcal{S}_\lambda)$ with a nonhomogeneous term. We construct its Green function, discuss its properties, and characterize the $L^2_{\mathcal{W}}$ solutions of this nonhomogeneous system in terms of the Green function (Theorem 8.5). A certain uniqueness result is also proven for the limit point case.

2. Time Scales

Following [52, 53], a time scale $\mathbb{T}$ is any nonempty and closed subset of $\mathbb{R}$. A bounded time scale can be therefore identified as $[a,b]_{\mathbb{T}} := [a,b] \cap \mathbb{T}$ which we call the time scale interval, where $a := \min \mathbb{T}$ and $b := \max \mathbb{T}$. Similarly, a time scale which is unbounded above has the form $[a,\infty)_{\mathbb{T}} := [a,\infty) \cap \mathbb{T}$. The forward and backward jump operators on a time scale are denoted by $\sigma(t)$ and $\rho(t)$ and the graininess function by $\mu(t) := \sigma(t) - t$. If not otherwise stated, all functions in this paper are considered to be complex valued. A function f on $[a,b]_{\mathbb{T}}$ is called *piecewise rd-continuous*; we write $f \in \mathrm{C}_{\mathrm{prd}}$ on $[a,b]_{\mathbb{T}}$ if the right-hand limit $f(t^+)$ exists finite at all right-dense points $t \in [a,b)_{\mathbb{T}}$, and the left-hand limit $f(t^-)$ exists finite at all left-dense points $t \in (a,b]_{\mathbb{T}}$ and f is continuous in the topology of the given time scale at all but possibly finitely many right-dense points $t \in [a,b)_{\mathbb{T}}$. A function f on $[a,\infty)_{\mathbb{T}}$ is *piecewise rd-continuous*; we write $f \in \mathrm{C}_{\mathrm{prd}}$ on $[a,\infty)_{\mathbb{T}}$ if $f \in \mathrm{C}_{\mathrm{prd}}$ on $[a,b]_{\mathbb{T}}$ for every $b \in (a,\infty)_{\mathbb{T}}$. An $n \times n$ matrix-valued function f is called *regressive* on a given time scale interval if $I + \mu(t) f(t)$ is invertible for all t in this interval.

The time scale Δ-derivative of a function f at a point t is denoted by $f^\Delta(t)$; see [52, Definition 1.10]. Whenever $f^\Delta(t)$ exists, the formula $f^\sigma(t) = f(t) + \mu(t) f^\Delta(t)$ holds true. The product rule for the Δ-differentiation of the product of two functions has the form

$$(fg)^\Delta(t) = f^\Delta(t)g(t) + f^\sigma(t)g^\Delta(t) = f^\Delta(t)g^\sigma(t) + f(t)g^\Delta(t). \tag{2.1}$$

A function f on $[a,b]_{\mathbb{T}}$ is called *piecewise rd-continuously Δ-differentiable*; we write $f \in \mathrm{C}^1_{\mathrm{prd}}$ on $[a,b]_{\mathbb{T}}$; if it is continuous on $[a,b]_{\mathbb{T}}$, then $f^\Delta(t)$ exists at all except for possibly finitely many points $t \in [a,\rho(b)]_{\mathbb{T}}$, and $f^\Delta \in \mathrm{C}_{\mathrm{prd}}$ on $[a,\rho(b)]_{\mathbb{T}}$. As a consequence we have that the finitely many points t_i at which $f^\Delta(t_i)$ does not exist belong to $(a,b)_{\mathbb{T}}$ and these points t_i are necessarily right-dense and left-dense at the same time. Also, since at those points we know that $f^\Delta(t_i^+)$ and $f^\Delta(t_i^-)$ exist finite, we replace the quantity $f^\Delta(t_i)$ by $f^\Delta(t_i^\pm)$ in any formula involving $f^\Delta(t)$ for all $t \in [a,\rho(b)]_{\mathbb{T}}$. Similarly as above we define $f \in \mathrm{C}^1_{\mathrm{prd}}$ on $[a,\infty)_{\mathbb{T}}$. The time scale integral of a piecewise rd-continuous function f over $[a,b]_{\mathbb{T}}$ is denoted by $\int_a^b f(t)\Delta t$ and over $[a,\infty)_{\mathbb{T}}$ by $\int_a^\infty f(t)\Delta t$ provided this integral is convergent in the usual sense; see [52, Definitions 1.71 and 1.82].

to Contents, page 1

Remark 2.1. As it is known in [52, Theorem 5.8] and discussed in [54, Remark 3.8], for a fixed $t_0 \in [a,b]_{\mathbb{T}}$ and a piecewise rd-continuous $n \times n$ matrix function $A(\cdot)$ on $[a,b]_{\mathbb{T}}$ which is regressive on $[a,t_0)_{\mathbb{T}}$, the initial value problem $y^{\Delta}(t) = A(t)y(t)$ for $t \in [a,\rho(b)]_{\mathbb{T}}$ with $y(t_0) = y_0$ has a unique solution $y(\cdot) \in \mathrm{C}^1_{\mathrm{prd}}$ on $[a,b]_{\mathbb{T}}$ for any $y_0 \in \mathbb{C}^n$. Similarly, this result holds on $[a,\infty)_{\mathbb{T}}$.

Let us recall some matrix notations from linear algebra used in this paper. Given a complex square matrix M, by M^*, $M > 0$, $M \geq 0$, $M < 0$, $M \leq 0$, $\operatorname{rank} M$, $\operatorname{Ker} M$, $\operatorname{def} M$, we denote, respectively, the conjugate transpose, positive definiteness, positive semidefiniteness, negative definiteness, negative semidefiniteness, rank, kernel, and the defect (i.e., the dimension of the kernel) of the matrix M. Moreover, we will use the notation $\operatorname{Im}(M) := (M - M^*)/(2i)$ and $\operatorname{Re}(M) := (M + M^*)/2$ for the Hermitian components of the matrix M; see [55, pages 268-269] or [56, Fact 3.5.24]. This notation will be also used with $\lambda \in \mathbb{C}$, and in this case $\operatorname{Im}(\lambda)$ and $\operatorname{Re}(\lambda)$ represent the imaginary and real parts of λ.

Remark 2.2. If the matrix $\operatorname{Im}(M)$ is positive or negative definite, then the matrix M is necessarily invertible. The proof of this fact can be found, for example, in [2, Remark 2.6].

In order to simplify the notation we abbreviate $[f^{\sigma}(t)]^*$ and $[f^*(t)]^{\sigma}$ by $f^{\sigma*}(t)$. Similarly, instead of $[f^{\Delta}(t)]^*$ and $[f^*(t)]^{\Delta}$ we will use $f^{\Delta*}(t)$.

3. Time Scale Symplectic Systems

Let $\mathcal{A}(\cdot), \mathcal{B}(\cdot), \mathcal{C}(\cdot), \mathfrak{D}(\cdot), \mathcal{W}(\cdot)$ be $n \times n$ piecewise rd-continuous functions on $[a,\infty)_{\mathbb{T}}$ such that $\mathcal{W}(t) \geq 0$ for all $t \in [a,\infty)_{\mathbb{T}}$; that is, $\mathcal{W}(t)$ is Hermitian and nonnegative definite, satisfying identity (1.8). In this paper we consider the linear system $(\mathcal{S}_\lambda)$ introduced in the previous section. This system can be written as

$$z^{\Delta}(t,\lambda) = \mathcal{S}(t)z(t,\lambda) + \lambda\mathcal{J}\widetilde{\mathcal{W}}(t)z^{\sigma}(t,\lambda), \quad t \in [a,\infty)_{\mathbb{T}}, \tag{$\mathcal{S}_\lambda$}$$

where the $2n \times 2n$ matrix $\widetilde{\mathcal{W}}(t)$ is defined and has the property

$$\widetilde{\mathcal{W}}(t) := \begin{pmatrix} \mathcal{W}(t) & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathcal{J}\widetilde{\mathcal{W}}(t) = \begin{pmatrix} 0 & 0 \\ -\mathcal{W}(t) & 0 \end{pmatrix}. \tag{3.1}$$

The system $(\mathcal{S}_\lambda)$ can be written in the equivalent form

$$z^{\Delta}(t,\lambda) = \mathcal{S}(t,\lambda)z(t,\lambda), \quad t \in [a,\infty)_{\mathbb{T}}, \tag{3.2}$$

to Contents, page 1

where the matrix $\mathcal{S}(t,\lambda)$ is defined through the matrices $\mathcal{S}(t)$ and $\widetilde{\mathcal{W}}(t)$ from (1.8) and (3.1) by

$$\begin{aligned}\mathcal{S}(t,\lambda) &:= \mathcal{S}(t) + \lambda\mathcal{J}\widetilde{\mathcal{W}}(t)\left[I+\mu(t)\mathcal{S}(t)\right]\\ &= \begin{pmatrix} \mathcal{A}(t) & \mathcal{B}(t) \\ \mathcal{C}(t) - \lambda\mathcal{W}(t)\left[I+\mu(t)\mathcal{A}(t)\right] & \mathcal{D}(t) - \lambda\mu(t)\mathcal{W}(t)\mathcal{B}(t)\end{pmatrix}.\end{aligned} \tag{3.3}$$

By using the identity in (1.8), a direct calculation shows that the matrix function $\mathcal{S}(\cdot,\cdot)$ satisfies

$$\mathcal{S}^*(t,\lambda)\mathcal{J} + \mathcal{J}\mathcal{S}\left(t,\overline{\lambda}\right) + \mu(t)\mathcal{S}^*(t,\lambda)\mathcal{J}\mathcal{S}\left(t,\overline{\lambda}\right) = 0, \quad t\in[a,\infty)_{\mathbb{T}},\ \lambda\in\mathbb{C}. \tag{3.4}$$

Here $\mathcal{S}^*(t,\lambda) = [\mathcal{S}(t,\lambda)]^*$, and $\overline{\lambda}$ is the usual conjugate number to λ.

Remark 3.1. The name time scale *symplectic system* or *Hamiltonian system* has been reserved in the literature for the system of the form

$$z^{\Delta}(t) = \mathbb{S}(t)z(t), \quad t\in[a,\infty)_{\mathbb{T}}, \tag{3.5}$$

in which the matrix function $\mathbb{S}(\cdot)$ satisfies the identity in (1.8); see [44–47, 57], and compare also, for example, with [58–61]. Since for a fixed $\lambda,\nu\in\mathbb{C}$ the matrix $\mathcal{S}(t,\lambda)$ from (3.3) satisfies

$$\mathcal{S}^*(t,\lambda)\mathcal{J} + \mathcal{J}\mathcal{S}(t,\nu) + \mu(t)\mathcal{S}^*(t,\lambda)\mathcal{J}\mathcal{S}(t,\nu) = \left(\overline{\lambda}-\nu\right)\left[I+\mu(t)\mathcal{S}^*(t)\right]\widetilde{\mathcal{W}}(t)\left[I+\mu(t)\mathcal{S}(t)\right], \tag{3.6}$$

it follows that the system $(\mathcal{S}_\lambda)$ is a true time scale symplectic system according to the above terminology only for $\lambda\in\mathbb{R}$, while strictly speaking $(\mathcal{S}_\lambda)$ *is not* a time scale symplectic system for $\lambda\in\mathbb{C}\setminus\mathbb{R}$. However, since $(\mathcal{S}_\lambda)$ is a perturbation of the time scale symplectic system $(\mathcal{S}_0)$ and since the important properties of time scale symplectic systems needed in the presented Weyl-Titchmarsh theory, such as (3.4) or (3.8), are satisfied in an appropriate modification, we accept with the above understanding the same terminology for the system $(\mathcal{S}_\lambda)$ for any $\lambda\in\mathbb{C}$.

Equation (3.4) represents a fundamental identity for the theory of time scale symplectic systems $(\mathcal{S}_\lambda)$. Some important properties of the matrix $\mathcal{S}(t,\lambda)$ are displayed below. Note that formula (3.7) is a generalization of [46, equation (10.4)] to complex values of λ.

Lemma 3.2. *Identity* (3.4) *is equivalent to the identity*

$$\mathcal{S}\left(t,\overline{\lambda}\right)\mathcal{J} + \mathcal{J}\mathcal{S}^*(t,\lambda) + \mu(t)\mathcal{S}\left(t,\overline{\lambda}\right)\mathcal{J}\mathcal{S}^*(t,\lambda) = 0, \quad t\in[a,\infty)_{\mathbb{T}},\ \lambda\in\mathbb{C}. \tag{3.7}$$

In this case for any $\lambda \in \mathbb{C}$ *we have*

$$[I + \mu(t)\mathcal{S}^*(t,\lambda)]\mathcal{J}\left[I + \mu(t)\mathcal{S}\left(t,\overline{\lambda}\right)\right] = \mathcal{J}, \quad t \in [a,\infty)_{\mathbb{T}}, \tag{3.8}$$

$$\left[I + \mu(t)\mathcal{S}\left(t,\overline{\lambda}\right)\right]\mathcal{J}[I + \mu(t)\mathcal{S}^*(t,\lambda)] = \mathcal{J}, \quad t \in [a,\infty)_{\mathbb{T}}, \tag{3.9}$$

and the matrices $I + \mu(t)\mathcal{S}(t,\lambda)$ *and* $I + \mu(t)\mathcal{S}(t,\overline{\lambda})$ *are invertible with*

$$[I + \mu(t)\mathcal{S}(t,\lambda)]^{-1} = -\mathcal{J}\left[I + \mu(t)\mathcal{S}^*\left(t,\overline{\lambda}\right)\right]\mathcal{J}, \quad t \in [a,\infty)_{\mathbb{T}}. \tag{3.10}$$

Proof. Let $t \in [a,\infty)_{\mathbb{T}}$ and $\lambda \in \mathbb{C}$ be fixed. If t is right-dense, that is, $\mu(t) = 0$, then identity (3.4) reduces to $\mathcal{S}^*(t,\lambda)\mathcal{J} + \mathcal{J}\mathcal{S}(t,\overline{\lambda}) = 0$. Upon multiplying this equation by $\mathcal{J}$ from the left and right side, we get identity (3.7) with $\mu(t) = 0$. If t is right scattered, that is, $\mu(t) > 0$, then (3.4) is equivalent to (3.8). It follows that the determinants of $I + \mu(t)\mathcal{S}(t,\lambda)$ and $I + \mu(t)\mathcal{S}(t,\overline{\lambda})$ are nonzero proving that these matrices are invertible with the inverse given by (3.10). Upon multiplying (3.8) by the invertible matrices $[I + \mu(t)\mathcal{S}(t,\overline{\lambda})]\mathcal{J}$ from the left and $-[I + \mu(t)\mathcal{S}(t,\overline{\lambda})]^{-1}\mathcal{J}$ from the right and by using $\mathcal{J}^2 = -I$, we get formula (3.9), which is equivalent to (3.7) due to $\mu(t) > 0$. □

Remark 3.3. Equation (3.10) allows writing the system $(\mathcal{S}_\lambda)$ in the equivalent adjoint form

$$z^{\Delta}(t,\lambda) = \mathcal{J}\mathcal{S}^*\left(t,\overline{\lambda}\right)\mathcal{J}z^{\sigma}(t,\lambda), \quad t \in [a,\infty)_{\mathbb{T}}. \tag{3.11}$$

System (3.11) can be found, for example, in [47, Remark 3.1(iii)] or [50, equation (3.2)] in the connection with optimality conditions for variational problems over time scales.

In the following result we show that (3.4) guarantees, among other properties, the existence and uniqueness of solutions of the initial value problems associated with $(\mathcal{S}_\lambda)$.

Theorem 3.4 (existence and uniqueness theorem). *Let* $\lambda \in \mathbb{C}$, $t_0 \in [a,\infty)_{\mathbb{T}}$, *and* $z_0 \in \mathbb{C}^{2n}$ *be given. Then the initial value problem* $(\mathcal{S}_\lambda)$ *with* $z(t_0) = z_0$ *has a unique solution* $z(\cdot,\lambda) \in \mathrm{C}^1_{\mathrm{prd}}$ *on the interval* $[a,\infty)_{\mathbb{T}}$.

Proof. The coefficient matrix of system $(\mathcal{S}_\lambda)$, or equivalently of system (3.2), is piecewise rd-continuous on $[a,\infty)_{\mathbb{T}}$. By Lemma 3.2, the matrix $I+\mu(t)\mathcal{S}(t,\lambda)$ is invertible for all $t \in [a,\infty)_{\mathbb{T}}$, which proves that the function $\mathcal{S}(\cdot,\lambda)$ is regressive on $[a,\infty)_{\mathbb{T}}$. Hence, the result follows from Remark 2.1. □

If not specified otherwise, we use a common agreement that $2n$-vector solutions of system $(\mathcal{S}_\lambda)$ and $2n \times n$-matrix solutions of system $(\mathcal{S}_\lambda)$ are denoted by small letters and capital letters, respectively, typically by $z(\cdot,\lambda)$ or $\tilde{z}(\cdot,\lambda)$ and $Z(\cdot,\lambda)$ or $\widetilde{Z}(\cdot,\lambda)$.

Next we establish several identities involving solutions of system $(\mathcal{S}_\lambda)$ or solutions of two such systems with different spectral parameters. The first result is the Lagrange identity known in the special cases of continuous time linear Hamiltonian systems in [11, Theorem 4.1] or [8, equation (2.23)], discrete linear Hamiltonian systems in [9, equation (2.55)] or [14, Lemma 2.2], discrete symplectic systems in [1, Lemma 2.6] or [2, Lemma 2.3], and time scale linear Hamiltonian systems in [3, Lemma 3.5] and [5, Theorem 2.2].

Theorem 3.5 (Lagrange identity). *Let $\lambda, \nu \in \mathbb{C}$ and $m \in \mathbb{N}$ be given. If $z(\cdot,\lambda)$ and $z(\cdot,\nu)$ are $2n\times m$ solutions of systems $(\mathcal{S}_\lambda)$ and $(\mathcal{S}_\nu)$, respectively, then*

$$[z^*(t,\lambda)\mathcal{J}z(t,\nu)]^\Delta = \left(\overline{\lambda}-\nu\right)z^{\sigma *}(t,\lambda)\widetilde{\mathcal{W}}(t)z^\sigma(t,\nu), \quad t \in [a,\infty)_\mathbb{T}. \tag{3.12}$$

Proof. Formula (3.12) follows from the time scales product rule (2.1) by using the relation $z^\sigma(t,\lambda) = [I+\mu(t)\mathcal{S}(t,\lambda)]z(t,\lambda)$ and identity (3.6). □

As consequences of Theorem 3.5, we obtain the following.

Corollary 3.6. *Let $\lambda, \nu \in \mathbb{C}$ and $m \in \mathbb{N}$ be given. If $z(\cdot,\lambda)$ and $z(\cdot,\nu)$ are $2n\times m$ solutions of systems $(\mathcal{S}_\lambda)$ and $(\mathcal{S}_\nu)$, respectively, then for all $t \in [a,\infty)_\mathbb{T}$ we have*

$$z^*(t,\lambda)\mathcal{J}z(t,\nu) = z^*(a,\lambda)\mathcal{J}z(a,\nu) + \left(\overline{\lambda}-\nu\right)\int_a^t z^{\sigma *}(s,\lambda)\widetilde{\mathcal{W}}(s)z^\sigma(s,\nu)\Delta s. \tag{3.13}$$

One can easily see that if $z(\cdot,\lambda)$ is a solution of system $(\mathcal{S}_\lambda)$, then $z(\cdot,\overline{\lambda})$ is a solution of system $(\mathcal{S}_{\overline{\lambda}})$. Therefore, Theorem 3.5 with $\nu=\overline{\lambda}$ yields a Wronskian-type property of solutions of system $(\mathcal{S}_\lambda)$.

Corollary 3.7. *Let $\lambda \in \mathbb{C}$ and $m \in \mathbb{N}$ be given. For any $2n\times m$ solution $z(\cdot,\lambda)$ of systems $(\mathcal{S}_\lambda)$*

$$z^*(t,\lambda)\mathcal{J}z\left(t,\overline{\lambda}\right) \equiv z^*(a,\lambda)\mathcal{J}z\left(a,\overline{\lambda}\right), \quad \text{is constant on } [a,\infty)_\mathbb{T}. \tag{3.14}$$

The following result gives another interesting property of solutions of system $(\mathcal{S}_\lambda)$ and $(\mathcal{S}_{\overline{\lambda}})$.

Lemma 3.8. *Let $\lambda \in \mathbb{C}$ and $m \in \mathbb{N}$ be given. For any $2n\times m$ solutions $z(\cdot,\lambda)$ and $\tilde{z}(\cdot,\lambda)$ of system $(\mathcal{S}_\lambda)$, the $2n\times 2n$ matrix function $K(\cdot,\lambda)$ defined by*

$$K(t,\lambda) := z(t,\lambda)\tilde{z}^*\left(t,\overline{\lambda}\right) - \tilde{z}(t,\lambda)z^*\left(t,\overline{\lambda}\right), \quad t \in [a,\infty)_\mathbb{T}, \tag{3.15}$$

satisfies the dynamic equation

$$K^\Delta(t,\lambda) = \mathcal{S}(t,\lambda)K(t,\lambda) + [I+\mu(t)\mathcal{S}(t,\lambda)]K(t,\lambda)\mathcal{S}^*\left(t,\overline{\lambda}\right), \quad t \in [a,\infty)_\mathbb{T}, \tag{3.16}$$

and the identities $K^*(t,\lambda) = -K(t,\overline{\lambda})$ and

$$K^\sigma(t,\lambda) = [I+\mu(t)\mathcal{S}(t,\lambda)]K(t,\lambda)\left[I+\mu(t)\mathcal{S}^*\left(t,\overline{\lambda}\right)\right], \quad t \in [a,\infty)_\mathbb{T}. \tag{3.17}$$

to Contents, page 1

Proof. Having that $z(\cdot,\lambda)$ and $\tilde{z}(\cdot,\lambda)$ are solutions of system $(\mathcal{S}_\lambda)$, it follows that $z(\cdot,\overline{\lambda})$ and $\tilde{z}(\cdot,\overline{\lambda})$ are solutions of system $(\mathcal{S}_{\overline{\lambda}})$. The results then follow by direct calculations. □

Remark 3.9. The content of Lemma 3.8 appears to be new both in the continuous and discrete time cases. Moreover, when the matrix function $K(\cdot,\lambda) \equiv K(\lambda)$ is constant, identity (3.17) yields for any right-scattered $t \in [a,\infty)_{\mathbb{T}}$ that

$$\mathcal{S}(t,\lambda)K(\lambda) + K(\lambda)\mathcal{S}^*\left(t,\overline{\lambda}\right) + \mu(t)\mathcal{S}(t,\lambda)K(\lambda)\mathcal{S}^*\left(t,\overline{\lambda}\right) = 0. \tag{3.18}$$

It is interesting to note that this formula is very much like (3.7). More precisely, identity (3.7) is a consequence of (3.18) for the case of $K(\lambda) \equiv \mathcal{J}$.

Next we present properties of certain fundamental matrices $\Psi(\cdot,\lambda)$ of system $(\mathcal{S}_\lambda)$, which are generalizations of the corresponding results in [46, Section 10.2] to complex λ. Some of these results can be proven under the weaker condition that the initial value of $\Psi(a,\lambda)$ does depend on λ and satisfies $\Psi^*(a,\lambda)\mathcal{J}\Psi(a,\overline{\lambda}) = \mathcal{J}$. However, these more general results will not be needed in this paper.

Lemma 3.10. *Let $\lambda \in \mathbb{C}$ be fixed. If $\Psi(\cdot,\lambda)$ is a fundamental matrix of system $(\mathcal{S}_\lambda)$ such that $\Psi(a,\lambda)$ is symplectic and independent of λ, then for any $t \in [a,\infty)_{\mathbb{T}}$ we have*

$$\Psi^*(t,\lambda)\mathcal{J}\Psi\left(t,\overline{\lambda}\right) = \mathcal{J}, \qquad \Psi^{-1}(t,\lambda) = -\mathcal{J}\Psi^*\left(t,\overline{\lambda}\right)\mathcal{J}, \qquad \Psi(t,\lambda)\mathcal{J}\Psi^*\left(t,\overline{\lambda}\right) = \mathcal{J}. \tag{3.19}$$

Proof. Identity (3.19)(i) is a consequence of Corollary 3.7, in which we use the fact that $\Psi(a,\lambda)$ is symplectic and independent of λ. The second identity in (3.19) follows from the first one, while the third identity is obtained from the equation $\Psi(t,\lambda)\Psi^{-1}(t,\lambda) = I$. □

Remark 3.11. If the fundamental matrix $\Psi(\cdot,\lambda) = (Z(\cdot,\lambda)\ \ \tilde{Z}(\cdot,\lambda))$ in Lemma 3.10 is partitioned into two $2n \times n$ blocks, then (3.19)(i) and (3.19)(iii) have, respectively, the form

$$Z^*(t,\lambda)\mathcal{J}Z\left(t,\overline{\lambda}\right) = 0, \qquad Z^*(t,\lambda)\mathcal{J}\tilde{Z}\left(t,\overline{\lambda}\right) = I, \qquad \tilde{Z}^*(t,\lambda)\mathcal{J}\tilde{Z}\left(t,\overline{\lambda}\right) = 0, \tag{3.20}$$

$$Z(t,\lambda)\tilde{Z}^*\left(t,\overline{\lambda}\right) - \tilde{Z}(t,\lambda)Z^*\left(t,\overline{\lambda}\right) = \mathcal{J}. \tag{3.21}$$

Observe that the matrix on the left-hand side of (3.21) represents a constant matrix $K(t,\lambda)$ from Lemma 3.8 and Remark 3.9.

Corollary 3.12. *Under the conditions of Lemma 3.10, for any $t \in [a,\infty)_{\mathbb{T}}$, we have*

$$\Psi^\sigma(t,\lambda)\mathcal{J}\Psi^*\left(t,\overline{\lambda}\right) = [I + \mu(t)\mathcal{S}(t,\lambda)]\mathcal{J}, \tag{3.22}$$

which in the notation of Remark 3.11 has the form

$$Z^\sigma(t,\lambda)\tilde{Z}^*\left(t,\overline{\lambda}\right) - \tilde{Z}^\sigma(t,\lambda)Z^*\left(t,\overline{\lambda}\right) = [I + \mu(t)\mathcal{S}(t,\lambda)]\mathcal{J}. \tag{3.23}$$

to Contents, page 1

Proof. Identity (3.22) follows from the equation $\Psi^{\sigma}(t,\lambda) = [I + \mu(t)\mathcal{S}(t,\lambda)]\Psi(t,\lambda)$ by applying formula (3.19)(ii). □

4. $M(\lambda)$-Function for Regular Spectral Problem

In this section we consider the regular spectral problem on the time scale interval $[a,b]_{\mathbb{T}}$ with some fixed $b \in (a,\infty)_{\mathbb{T}}$. We will specify the corresponding boundary conditions in terms of complex $n \times 2n$ matrices from the set

$$\Gamma := \left\{\alpha \in \mathbb{C}^{n\times 2n},\ \alpha\alpha^* = I,\ \alpha\mathcal{J}\alpha^* = 0\right\}. \tag{4.1}$$

The two defining conditions for $\alpha \in \mathbb{C}^{n\times 2n}$ in (4.1) imply that the $2n \times 2n$ matrix $(\alpha^* \quad -\mathcal{J}\alpha^*)$ is unitary and symplectic. This yields the identity

$$(\alpha^* \quad -\mathcal{J}\alpha^*)\begin{pmatrix}\alpha \\ \alpha\mathcal{J}\end{pmatrix} = I \in \mathbb{C}^{2n\times 2n}, \quad \text{that is,}\ \ \alpha^*\alpha - \mathcal{J}\alpha^*\alpha\mathcal{J} = I. \tag{4.2}$$

The last equation also implies, compare with [60, Remark 2.1.2], that

$$\operatorname{Ker}\alpha = \operatorname{Im}\mathcal{J}\alpha^*. \tag{4.3}$$

Let $\alpha, \beta \in \Gamma$ be fixed and consider the boundary value problem

$$(\mathcal{S}_\lambda), \quad \alpha z(a,\lambda) = 0, \quad \beta z(b,\lambda) = 0. \tag{4.4}$$

Our first result shows that the boundary conditions in (4.4) are equivalent with the boundary conditions phrased in terms of the images of the $2n \times 2n$ matrices

$$R_a := (\mathcal{J}\alpha^* \quad 0), \qquad R_b := (0 \quad -\mathcal{J}\beta^*), \tag{4.5}$$

which satisfy $R_a^*\mathcal{J}R_a = 0$, $R_b^*\mathcal{J}R_b = 0$, and $\operatorname{rank}(R_a^* \quad R_b^*) = 2n$.

Lemma 4.1. *Let $\alpha, \beta \in \Gamma$ and $\lambda \in \mathbb{C}$ be fixed. A solution $z(\cdot,\lambda)$ of system $(\mathcal{S}_\lambda)$ satisfies the boundary conditions in* (4.4) *if and only if there exists a unique vector $\xi \in \mathbb{C}^{2n}$ such that*

$$z(a,\lambda) = R_a\xi, \qquad z(b,\lambda) = R_b\xi. \tag{4.6}$$

Proof. Assume that (4.4) holds. Identity (4.3) implies the existence of vectors $\xi_a, \xi_b \in \mathbb{C}^n$ such that $z(a,\lambda) = -\mathcal{J}\alpha^*\xi_a$ and $z(b,\lambda) = -\mathcal{J}\beta^*\xi_b$. It follows that $z(\cdot,\lambda)$ satisfies (4.6) with $\xi := (-\xi_a^* \quad \xi_b^*)^*$. It remains to prove that ξ is unique such a vector. If $z(\cdot,\lambda)$ satisfies (4.6) and also $z(a,\lambda) = R_a\zeta$ and $z(b,\lambda) = R_b\zeta$ for some $\xi, \zeta \in \mathbb{C}^{2n}$, then $R_a(\xi-\zeta) = 0$ and $R_b(\xi-\zeta) = 0$. Hence, $\mathcal{J}\alpha^*(I \quad 0)(\xi-\zeta) = 0$ and $-\mathcal{J}\beta^*(0 \quad I)(\xi-\zeta) = 0$. If we multiply the latter two equalities by $\alpha\mathcal{J}$ and $\beta\mathcal{J}$, respectively, and use $\alpha\alpha^* = I = \beta\beta^*$, then we obtain $(I \quad 0)(\xi - \zeta) = 0$ and $(0 \quad I)(\xi - \zeta) = 0$.

to Contents, page 1

This yields $\xi - \zeta = 0$, which shows that the vector ξ in (4.6) is unique. The opposite direction, that is, that (4.6) implies (4.4), is trivial. □

Following the standard terminology, see, for example, [62, 63], a number $\lambda \in \mathbb{C}$ is an *eigenvalue* of (4.4) if this boundary value problem has a solution $z(\cdot,\lambda) \not\equiv 0$. In this case the function $z(\cdot,\lambda)$ is called the *eigenfunction* corresponding to the eigenvalue λ, and the dimension of the space of all eigenfunctions corresponding to λ (together with the zero function) is called the *geometric multiplicity* of λ.

Given $\alpha \in \Gamma$, we will utilize from now on the fundamental matrix $\Psi(\cdot,\lambda,\alpha)$ of system $(\mathcal{S}_\lambda)$ satisfying the initial condition from (4.4), that is,

$$\Psi^{\Delta}(t,\lambda,\alpha) = \mathcal{S}(t,\lambda)\Psi(t,\lambda,\alpha), \quad t \in [a,\rho(b)]_{\mathbb{T}}, \quad \Psi(a,\lambda,\alpha) = \begin{pmatrix}\alpha^* & -\mathcal{J}\alpha^*\end{pmatrix}. \tag{4.7}$$

Then $\Psi(a,\lambda,\alpha)$ does not depend on λ, and it is symplectic and unitary with the inverse $\Psi^{-1}(a,\lambda,\alpha) = \Psi^*(a,\lambda,\alpha)$. Hence, the properties of fundamental matrices derived earlier in Lemma 3.10, Remark 3.11, and Corollary 3.12 apply for the matrix function $\Psi(\cdot,\lambda,\alpha)$.

The following assumption will be imposed in this section when studying the regular spectral problem.

Hypothesis 4.2. For every $\lambda \in \mathbb{C}$, we have

$$\int_a^b \Psi^{\sigma *}(t,\lambda,\alpha)\widetilde{\mathcal{W}}(t)\Psi^{\sigma}(t,\lambda,\alpha)\Delta t > 0. \tag{4.8}$$

Condition (4.8) can be written in the equivalent form as

$$\int_a^b z^{\sigma *}(t,\lambda)\widetilde{\mathcal{W}}(t)z^{\sigma}(t,\lambda)\Delta t > 0, \tag{4.9}$$

for every nontrivial solution $z(\cdot,\lambda)$ of system $(\mathcal{S}_\lambda)$. Assumptions (4.8) and (4.9) are equivalent by a simple argument using the uniqueness of solutions of system $(\mathcal{S}_\lambda)$. The latter form (4.9) has been widely used in the literature, such as in the continuous time case in [8, Hypothesis 2.2], [30, equation (1.3)], [26, equation (2.3)], in the discrete time case in [9, Condition (2.16)], [14, equation (1.7)], [1, Assumption 2.2], [2, Hypothesis 2.4], and in the time scale Hamiltonian case in [3, Assumption 3] and [5, Condition (3.9)].

Following Remark 3.11, we partition the fundamental matrix $\Psi(\cdot,\lambda,\alpha)$ as

$$\Psi(\cdot,\lambda,\alpha) = \begin{pmatrix}Z(\cdot,\lambda,\alpha) & \widetilde{Z}(\cdot,\lambda,\alpha)\end{pmatrix}, \tag{4.10}$$

where $Z(\cdot,\lambda,\alpha)$ and $\widetilde{Z}(\cdot,\lambda,\alpha)$ are the $2n \times n$ solutions of system $(\mathcal{S}_\lambda)$ satisfying $Z(a,\lambda,\alpha) = \alpha^*$ and $\widetilde{Z}(a,\lambda,\alpha) = -\mathcal{J}\alpha^*$. With the notation

$$\Lambda(\lambda,\alpha,\beta) := \Psi(b,\lambda,\alpha)\Psi^*(a,\lambda,\alpha)R_a - R_b = \begin{pmatrix}-\widetilde{Z}(b,\lambda,\alpha) & \mathcal{J}\beta^*\end{pmatrix}, \tag{4.11}$$

to Contents, page 1

we have the classical characterization of the eigenvalues of (4.4); see, for example, the continuous time in [64, Chapter 4], the discrete time in [14, Theorem 2.3, Lemma 2.4], [2, Lemma 2.9, Theorem 2.11], and the time scale case in [62, Lemma 3], [63, Corollary 1].

Proposition 4.3. *For $\alpha, \beta \in \Gamma$ and $\lambda \in \mathbb{C}$, we have with notation* (4.11) *the following.*

(i) *The number λ is an eigenvalue of* (4.4) *if and only if* $\det \Lambda(\lambda, \alpha, \beta) = 0$.

(ii) *The algebraic multiplicity of the eigenvalue λ, that is, the number* $\operatorname{def} \Lambda(\lambda, \alpha, \beta)$, *is equal to the geometric multiplicity of λ.*

(iii) *Under Hypothesis 4.2, the eigenvalues of* (4.4) *are real, and the eigenfunctions corresponding to different eigenvalues are orthogonal with respect to the semi-inner product*

$$\langle z(\cdot, \lambda), z(\cdot, \nu) \rangle_{\mathcal{W},b} := \int_a^b z^{\sigma *}(t, \lambda) \widetilde{\mathcal{W}}(t) z^{\sigma}(t, \nu) \Delta t. \tag{4.12}$$

Proof. The arguments are here standard, and we refer to [44, Section 5], [63, Corollary 1], [3, Theorem 3.6]. □

The next algebraic characterization of the eigenvalues of (4.4) is more appropriate for the development of the Weyl-Titchmarsh theory for (4.4), since it uses the matrix $\beta \tilde{Z}(b, \lambda, \alpha)$ which has dimension n instead of using the matrix $\Lambda(\lambda, \alpha, \beta)$ which has dimension $2n$. Results of this type can be found in special cases of system $(\mathcal{S}_\lambda)$ in [8, Lemma 2.5], [11, Theorem 4.1], [9, Lemma 2.8], [14, Lemma 3.1], [1, Lemma 2.5], [3, Theorem 3.4], and [2, Lemma 3.1].

Lemma 4.4. *Let $\alpha, \beta \in \Gamma$ and $\lambda \in \mathbb{C}$ be fixed. Then λ is an eigenvalue of* (4.4) *if and only if* $\det \beta \tilde{Z}(b, \lambda, \alpha) = 0$. *In this case the algebraic and geometric multiplicities of λ are equal to* $\operatorname{def} \beta \tilde{Z}(b, \lambda, \alpha)$.

Proof. One can follow the same arguments as in the proof of the corresponding discrete symplectic case in [2, Lemma 3.1]. However, having the result of Proposition 4.3, we can proceed directly by the methods of linear algebra. In this proof we abbreviate $\Lambda := \Lambda(\lambda, \alpha, \beta)$ and $\tilde{Z} := \tilde{Z}(b, \lambda, \alpha)$. Assume that Λ is singular, that is, $-\tilde{Z}c + \mathcal{J}\beta^* d = 0$ for some vectors $c, d \in \mathbb{C}^n$, not both zero. Then $\tilde{Z}c = \mathcal{J}\beta^* d$, which yields that $\beta\tilde{Z}c = 0$. If $c = 0$, then $\mathcal{J}\beta^* d = 0$, which implies upon the multiplication by $\beta\mathcal{J}$ from the left that $d = 0$. Since not both c and d can be zero, it follows that $c \neq 0$ and the matrix $\beta\tilde{Z}$ is singular. Conversely, if $\beta\tilde{Z}c = 0$ for some nonzero vector $c \in \mathbb{C}^n$, then $-\tilde{Z}c + \mathcal{J}\beta^* d = 0$; that is, Λ is singular, with the vector $d := -\beta\mathcal{J}\tilde{Z}c$. Indeed, by using identity (4.2) we have $\mathcal{J}\beta^* d = -\mathcal{J}\beta^*\beta\mathcal{J}\tilde{Z}c = (I - \beta^*\beta)\tilde{Z}c = \tilde{Z}c$. From the above we can also see that the number of linearly independent vectors in $\operatorname{Ker} \beta\tilde{Z}$ is the same as the number of linearly independent vectors in $\operatorname{Ker} \Lambda$. Therefore, by Proposition 4.3(ii), the algebraic and geometric multiplicities of λ as an eigenvalue of (4.4) are equal to $\operatorname{def} \beta\tilde{Z}$. □

Since the eigenvalues of (4.4) are real, it follows that the matrix $\beta\tilde{Z}(b, \lambda, \alpha)$ is invertible for every $\lambda \in \mathbb{C}$ except for at most n real numbers. This motivates the definition of the $M(\lambda)$-function for the regular spectral problem.

to Contents, page 1

Definition 4.5 ($M(\lambda)$-function). Let $\alpha, \beta \in \Gamma$. Whenever the matrix $\beta\widetilde{Z}(b,\lambda,\alpha)$ is invertible for some value $\lambda \in \mathbb{C}$, we define the *Weyl-Titchmarsh $M(\lambda)$-function* as the $n \times n$ matrix

$$M(\lambda) = M(\lambda, b) = M(\lambda, b, \alpha, \beta) := -\left[\beta\widetilde{Z}(b,\lambda,\alpha)\right]^{-1}\beta Z(b,\lambda,\alpha). \tag{4.13}$$

The above definition of the $M(\lambda)$-function is a generalization of the corresponding definitions for the continuous and discrete linear Hamiltonian and symplectic systems in [8, Definition 2.6], [9, Definition 2.9], [14, equation (3.10)], [1, page 2859], [2, Definition 3.2] and time scale linear Hamiltonian systems in [3, equation (4.1)]. The dependence of the $M(\lambda)$-function on b, α, and β will be suppressed in the notation, and $M(\lambda, b)$ or $M(\lambda, b, \alpha, \beta)$ will be used only in few situations when we emphasize the dependence on b (such as at the end of Section 5) or on α and β (as in Lemma 4.14). By [65, Corollary 4.5], see also [44, Remark 2.2], the $M(\cdot)$-function is an entire function in λ. Another important property of the $M(\lambda)$-function is established in the following.

Lemma 4.6. *Let $\alpha, \beta \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Then*

$$M^*(\lambda) = M\left(\overline{\lambda}\right). \tag{4.14}$$

Proof. We abbreviate $Z(\lambda) := Z(b,\lambda,\alpha)$ and $\widetilde{Z}(\lambda) := \widetilde{Z}(b,\lambda,\alpha)$. By using the definition of $M(\lambda)$ in (4.13) and identity (3.21), we have

$$\begin{aligned} M^*(\lambda) - M\left(\overline{\lambda}\right) &= \left[\beta\widetilde{Z}\left(\overline{\lambda}\right)\right]^{-1}\beta\left[Z\left(\overline{\lambda}\right)\widetilde{Z}^*(\lambda) - \widetilde{Z}\left(\overline{\lambda}\right)Z^*(\lambda)\right]\beta^*\left[\beta\widetilde{Z}(\lambda)\right]^{*-1} \\ &\overset{(3.21)}{=} \left[\beta\widetilde{Z}\left(\overline{\lambda}\right)\right]^{-1}\beta\mathcal{J}\beta^*\left[\beta\widetilde{Z}(\lambda)\right]^{*-1} = 0, \end{aligned} \tag{4.15}$$

because $\beta \in \Gamma$. Hence, equality (4.14) holds true. □

The following solution plays an important role in particular in the results concerning the square integrable solutions of system $(\mathcal{S}_\lambda)$.

Definition 4.7 (Weyl solution). For any matrix $M \in \mathbb{C}^{n\times n}$, we define the so-called *Weyl solution* of system $(\mathcal{S}_\lambda)$ by

$$\mathcal{X}(\cdot,\lambda,\alpha,M) := \Psi(\cdot,\lambda,\alpha)\left(I \quad M^*\right)^* = Z(\cdot,\lambda,\alpha) + \widetilde{Z}(\cdot,\lambda,\alpha)M, \tag{4.16}$$

where $Z(\cdot,\lambda,\alpha)$ and $\widetilde{Z}(\cdot,\lambda,\alpha)$ are defined in (4.10).

The function $\mathcal{X}(\cdot,\lambda,\alpha,M)$, being a linear combination of two solutions of system $(\mathcal{S}_\lambda)$, is also a solution of this system. Moreover, $\alpha\mathcal{X}(a,\lambda,\alpha,M) = I$, and, if $\beta\widetilde{Z}(b,\lambda,\alpha)$ is invertible, then $\beta\widetilde{\mathcal{X}}(b,\lambda,\alpha,M) = \beta\widetilde{Z}(b,\lambda,\alpha)[M - M(\lambda)]$. Consequently, if we take $M := M(\lambda)$ in Definition 4.7, then $\beta\mathcal{X}(b,\lambda,\alpha,M(\lambda)) = 0$; that is, the Weyl solution $\mathcal{X}(\cdot,\lambda,\alpha,M(\lambda))$ satisfies the right endpoint boundary condition in (4.4).

to Contents, page 1

Following the corresponding notions in [8, equation (2.18)], [9, equation (2.51)], [14, page 471], [1, page 2859], [2, equation (3.13)], [3, equation (4.2)], we define the Hermitian $n \times n$ matrix function $\mathcal{E}(M)$ for system $(\mathcal{S}_\lambda)$.

Definition 4.8. For a fixed $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$, we define the matrix function

$$\mathcal{E} : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{n\times n}, \qquad \mathcal{E}(M) = \mathcal{E}(M,b) := i\delta(\lambda)\mathcal{X}^*(b,\lambda,\alpha,M)\mathcal{J}\mathcal{X}(b,\lambda,\alpha,M), \tag{4.17}$$

where $\delta(\lambda) := \operatorname{sgn}\operatorname{Im}(\lambda)$.

For brevity we suppress the dependence of the function $\mathcal{E}(\cdot)$ on b and λ. In few cases we will need $\mathcal{E}(M)$ depending on b (as in Theorem 5.1 and Definition 6.2) and in such situations we will use the notation $\mathcal{E}(M,b)$. Since $(i\mathcal{J})^* = i\mathcal{J}$, it follows that $\mathcal{E}(M)$ is a Hermitian matrix for any $M \in \mathbb{C}^{n\times n}$. Moreover, from Corollary 3.6, we obtain the identity

$$\mathcal{E}(M) = -2\delta(\lambda)\operatorname{Im}(M) + 2|\operatorname{Im}(\lambda)| \int_a^b \mathcal{X}^{\sigma *}(t,\lambda,\alpha,M)\widetilde{\mathcal{W}}(t)\mathcal{X}^{\sigma}(t,\lambda,\alpha,M)\Delta t, \tag{4.18}$$

where we used the fact that

$$\mathcal{X}^*(a,\lambda,\alpha,M)\mathcal{J}\mathcal{X}(a,\lambda,\alpha,M) \overset{(4.7)}{=} M - M^* = 2i\operatorname{Im}(M). \tag{4.19}$$

Next we define the Weyl disk and Weyl circle for the regular spectral problem. The geometric characterizations of the Weyl disk and Weyl circle in terms of the contractive or unitary matrices which justify the terminology "disk" or "circle" will be presented in Section 5.

Definition 4.9 (Weyl disk and Weyl circle). For a fixed $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$, the set

$$D(\lambda) = D(\lambda,b) := \{M \in \mathbb{C}^{n\times n},\ \mathcal{E}(M) \le 0\}, \tag{4.20}$$

is called the *Weyl disk*, and the set

$$C(\lambda) = C(\lambda,b) := \partial D(\lambda) = \{M \in \mathbb{C}^{n\times n},\ \mathcal{E}(M) = 0\}, \tag{4.21}$$

is called the *Weyl circle*.

The dependence of the Weyl disk and Weyl circle on b will be again suppressed. In the following result we show that the Weyl circle consists of precisely those matrices $M(\lambda)$ with $\beta \in \Gamma$. This result generalizes the corresponding statements in [8, Lemma 2.8], [9, Lemma 2.13], [14, Lemma 3.3], [1, Theorem 3.1], [2, Theorem 3.6], and [3, Theorem 4.2].

Theorem 4.10. *Let $\alpha \in \Gamma$, $\lambda \in \mathbb{C}\setminus\mathbb{R}$, and $M \in \mathbb{C}^{n\times n}$. The matrix M belongs to the Weyl circle $C(\lambda)$ if and only if there exists $\beta \in \Gamma$ such that $\beta\mathcal{X}(b,\lambda,\alpha,M) = 0$. In this case and under Hypothesis 4.2, we have with such a matrix β that $M = M(\lambda)$ as defined in* (4.13).

to Contents, page 1

Proof. Assume that $M \in C(\lambda)$, that is, $\mathcal{E}(M) = 0$. Then, with the vector

$$\beta := \mathcal{X}^*(b)\mathcal{J} = \begin{pmatrix} I & M^* \end{pmatrix} \Psi^*(b,\lambda,\alpha)\mathcal{J} \in \mathbb{C}^{n\times 2n}, \tag{4.22}$$

where $\mathcal{X}(b)$ denotes $\mathcal{X}(b,\lambda,\alpha,M)$, we have

$$\beta\mathcal{X}(b) = \mathcal{X}^*(b)\mathcal{J}\mathcal{X}(b) = \left[\frac{1}{(i\delta(\lambda))}\right]\mathcal{E}(M) = 0. \tag{4.23}$$

Moreover, $\operatorname{rank}\beta = n$, because the matrices $\Psi(b,\lambda,\alpha)$ and $\mathcal{J}$ are invertible and $\operatorname{rank}\begin{pmatrix} I & M^* \end{pmatrix} = n$. In addition, the identity $\mathcal{J}^* = \mathcal{J}^{-1}$ yields

$$\beta\mathcal{J}\beta^* = \mathcal{X}^*(b)\mathcal{J}\mathcal{X}(b) \overset{(4.23)}{=} 0. \tag{4.24}$$

Now, if the condition $\beta\beta^* = I$ is not satisfied, then we replace β by $\widetilde{\beta} := (\beta\beta^*)^{-1/2}\beta$ (note that $\beta\beta^* > 0$, so that $(\beta\beta^*)^{-1/2}$ is well defined), and in this case

$$\begin{aligned} \widetilde{\beta}\mathcal{X}(b) &= (\beta\beta^*)^{-1/2}\beta\mathcal{X}(b) \overset{(4.23)}{=} 0, \\ \widetilde{\beta}\mathcal{J}\widetilde{\beta}^* &= (\beta\beta^*)^{-1/2}\beta\mathcal{J}\beta^*(\beta\beta^*)^{-1/2} \overset{(4.24)}{=} 0, \\ \widetilde{\beta}\widetilde{\beta}^* &= (\beta\beta^*)^{-1/2}\beta\beta^*(\beta\beta^*)^{-1/2} = I. \end{aligned} \tag{4.25}$$

Conversely, suppose that for a given $M \in \mathbb{C}^{n\times n}$ there exists $\beta \in \Gamma$ such that $\beta\mathcal{X}(b) = 0$. Then from (4.3) it follows that $\mathcal{X}(b) = \mathcal{J}\beta^* P$ for the matrix $P := -\beta\mathcal{J}\mathcal{X}(b) \in \mathbb{C}^{n\times n}$. Hence,

$$\mathcal{E}(M) = i\delta(\lambda)P^*\beta\mathcal{J}^*\mathcal{J}\mathcal{J}\beta^*P = i\delta(\lambda)P^*\beta\mathcal{J}\beta^*P = 0, \tag{4.26}$$

that is, $M \in C(\lambda)$. Finally, since $\lambda \in \mathbb{C} \setminus \mathbb{R}$, then by Proposition 4.3(iii) the number λ is not an eigenvalue of (4.4), which by Lemma 4.4 shows that the matrix $\beta\widetilde{Z}(b,\lambda,\alpha)$ is invertible. The definition of the Weyl solution in (4.16) then yields

$$\beta Z(b,\lambda,\alpha) + \beta\widetilde{Z}(b,\lambda,\alpha)M = \beta\mathcal{X}(b,\lambda,\alpha,M) = 0, \tag{4.27}$$

which implies that $M = -\left[\beta\widetilde{Z}(b,\lambda,\alpha)\right]^{-1}\beta Z(b,\lambda,\alpha) = M(\lambda)$. □

Remark 4.11. The matrix $P := -\beta\mathcal{J}\mathcal{X}(b,\lambda,\alpha,M) \in \mathbb{C}^{n\times n}$ from the proof of Theorem 4.10 is invertible. This fact was not needed in that proof. However, we show that P is invertible because this argument will be used in the proof of Lemma 4.14. First we prove that $\operatorname{Ker} P = \operatorname{Ker}\mathcal{X}(b,\lambda,\alpha,M)$. For if $Pd = 0$ for some $d \in \mathbb{C}^n$, then from identity (4.2) we get $\mathcal{X}(b,\lambda,\alpha,M)d = (I - \beta^*\beta)\mathcal{X}(b,\lambda,\alpha,M)d = \mathcal{J}\beta^*Pd = 0$. Therefore, $\operatorname{Ker} P \subseteq \operatorname{Ker}\mathcal{X}(b,\lambda,\alpha,M)$. The opposite inclusion follows by the definition of P. And since, by (4.16), $\operatorname{rank}\mathcal{X}(b,\lambda,\alpha,M) = \operatorname{rank}\begin{pmatrix} I & M^* \end{pmatrix}^* = n$, it follows that $\operatorname{Ker}\mathcal{X}(b,\lambda,\alpha,M) = \{0\}$. Hence, $\operatorname{Ker} P = \{0\}$ as well; that is, the matrix P is invertible.

to Contents, page 1

The next result contains a characterization of the matrices $M \in \mathbb{C}^{n\times n}$ which lie "inside" the Weyl disk $D(\lambda)$. In the previous result (Theorem 4.10) we have characterized the elements of the boundary of the Weyl disk $D(\lambda)$, that is, the elements of the Weyl circle $C(\lambda)$, in terms of the matrices $\beta \in \Gamma$. For such β we have $\beta\mathcal{J}\beta^* = 0$, which yields $i\delta(\lambda)\beta\mathcal{J}\beta^* = 0$. Comparing with that statement we now utilize the matrices $\beta \in \mathbb{C}^{n\times 2n}$ which satisfy $i\delta(\lambda)\beta\mathcal{J}\beta^* > 0$. In the special cases of the continuous and discrete time, this result can be found in [8, Lemma 2.13], [9, Lemma 2.18], and [2, Theorem 3.13].

Theorem 4.12. *Let $\alpha \in \Gamma$, $\lambda \in \mathbb{C} \setminus \mathbb{R}$, and $M \in \mathbb{C}^{n\times n}$. The matrix M satisfies $\mathcal{E}(M) < 0$ if and only if there exists $\beta \in \mathbb{C}^{n\times 2n}$ such that $i\delta(\lambda)\beta\mathcal{J}\beta^* > 0$ and $\beta\mathcal{X}(b,\lambda,\alpha,M) = 0$. In this case and under Hypothesis 4.2, we have with such a matrix β that $M = M(\lambda)$ as defined in* (4.13) *and β may be chosen so that $\beta\beta^* = I$.*

Proof. For $M \in \mathbb{C}^{n\times n}$ consider on $[a,b]_{\mathbb{T}}$ the Weyl solution

$$\mathcal{X}(\cdot) := \mathcal{X}(\cdot,\lambda,\alpha,M) = \begin{pmatrix} \mathcal{X}_1(\cdot) \\ \mathcal{X}_2(\cdot) \end{pmatrix}, \quad \text{with } n\times n \text{ blocks } \mathcal{X}_1(\cdot) \text{ and } \mathcal{X}_2(\cdot). \tag{4.28}$$

Suppose first that $\mathcal{E}(M) < 0$. Then the matrices $\mathcal{X}_j(b)$, $j \in \{1,2\}$, are invertible. Indeed, if one of them is singular, then there exists a nonzero vector $v \in \mathbb{C}^n$ such that $\mathcal{X}_1(b)v = 0$ or $\mathcal{X}_2(b)v = 0$. Then

$$v^*\mathcal{E}(M)v = i\delta(\lambda)v^*\mathcal{X}^*(b)\mathcal{J}\mathcal{X}(b)v = i\delta(\lambda)v^*\left[\mathcal{X}_1^*(b)\mathcal{X}_2(b) - \mathcal{X}_2^*(b)\mathcal{X}_1(b)\right]v = 0, \tag{4.29}$$

which contradicts $\mathcal{E}(M) < 0$. Now we set $\beta_1 := I$, $\beta_2 := -\mathcal{X}_1(b)\mathcal{X}_2^{-1}(b)$, and $\beta := (\beta_1 \;\; \beta_2)$. Then for this $2n\times n$ matrix β we have $\beta\mathcal{X}(b) = 0$ and, by a similar calculation as in (4.29),

$$\begin{aligned} \mathcal{E}(M) &= i\delta(\lambda)\mathcal{X}^*(b)\mathcal{J}\mathcal{X}(b) = i\delta(\lambda)\mathcal{X}_2^*(b)\left(\beta_2\beta_1^* - \beta_1\beta_2^*\right)\mathcal{X}_2(b) \\ &= 2\delta(\lambda)\mathcal{X}_2^*(b)\operatorname{Im}\left(\beta_1\beta_2^*\right)\mathcal{X}_2(b) = -i\delta(\lambda)\mathcal{X}_2^*(b)\beta\mathcal{J}\beta^*\mathcal{X}_2(b), \end{aligned} \tag{4.30}$$

where we used the equality $\beta\mathcal{J}\beta^* = 2i\operatorname{Im}(\beta_1\beta_2^*)$. Since $\mathcal{E}(M) < 0$ and $\mathcal{X}_2(b)$ is invertible, it follows that $i\delta(\lambda)\beta\mathcal{J}\beta^* > 0$. Conversely, assume that for a given matrix $M \in \mathbb{C}^{n\times n}$ there is $\beta = (\beta_1 \;\; \beta_2) \in \mathbb{C}^{n\times 2n}$ satisfying $i\delta(\lambda)\beta\mathcal{J}\beta^* > 0$ and $\beta\mathcal{X}(b) = 0$. Condition $i\delta(\lambda)\beta\mathcal{J}\beta^* > 0$ is equivalent to $\operatorname{Im}(\beta_1\beta_2^*) < 0$ when $\operatorname{Im}(\lambda) > 0$ and to $\operatorname{Im}(\beta_1\beta_2^*) > 0$ when $\operatorname{Im}(\lambda) < 0$. The positive or negative definiteness of $\operatorname{Im}(\beta_1\beta_2^*)$ implies the invertibility of β_1 and β_2; see Remark 2.2. Therefore, from the equality $\beta_1\mathcal{X}_1(b) + \beta_2\mathcal{X}_2(b) = \beta\mathcal{X}(b) = 0$, we obtain $\mathcal{X}_1(b) = -\beta_1^{-1}\beta_2\mathcal{X}_2(b)$, and so

$$\begin{aligned} \mathcal{E}(M) &= i\delta(\lambda)\left[\mathcal{X}_1^*(b)\mathcal{X}_2(b) - \mathcal{X}_2^*(b)\mathcal{X}_1(b)\right] \\ &= i\delta(\lambda)\mathcal{X}_2^*(b)\beta_1^{-1}\left(\beta_2\beta_1^* - \beta_1\beta_2^*\right)\beta_1^{*-1}\mathcal{X}_2(b) \\ &= -i\delta(\lambda)\mathcal{X}_2^*(b)\beta_1^{-1}\beta\mathcal{J}\beta^*\beta_1^{*-1}\mathcal{X}_2(b). \end{aligned} \tag{4.31}$$

to Contents, page 1

The matrix $\mathcal{X}_2(b)$ is invertible, because if $\mathcal{X}_2(b)d = 0$ for some nonzero vector $d \in \mathbb{C}^n$, then $\mathcal{X}_1(b)d = -\beta_1^{-1}\beta_2\mathcal{X}_2(b)d = 0$, showing that $\operatorname{rank}\mathcal{X}(b) < n$. This however contradicts $\operatorname{rank}\mathcal{X}(b) = n$ which we have from the definition of the Weyl solution $\mathcal{X}(\cdot)$ in (4.16). Consequently, (4.31) yields through $i\delta(\lambda)\beta\mathcal{J}\beta^* > 0$ that $\mathcal{E}(M) < 0$.

If the matrix β does not satisfy $\beta\beta^* = I$, then we modify it according to the procedure described in the proof of Theorem 4.10. Finally, since $\lambda \in \mathbb{C} \setminus \mathbb{R}$, we get from Proposition 4.3(iii) and Lemma 4.4 that the matrix $\beta\widetilde{Z}(b,\lambda,\alpha)$ is invertible which in turn implies through the calculation in (4.27) that $M = -[\beta\widetilde{Z}(b,\lambda,\alpha)]^{-1}\beta Z(b,\lambda,\alpha) = M(\lambda)$. □

In the following lemma we derive some additional properties of the Weyl disk and the $M(\lambda)$-function. Special cases of this statement can be found in [8, Lemma 2.9], [33, Theorem 3.1], [9, Lemma 2.14], [14, Lemma 3.2(ii)], [1, Theorem 3.7], [2, Lemma 3.7], and [3, Theorem 4.13].

Theorem 4.13. *Let $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. For any matrix $M \in D(\lambda)$ we have*

$$\delta(\lambda)\operatorname{Im}(M) \geq |\operatorname{Im}(\lambda)| \int_a^b \mathcal{X}^{\sigma *}(t,\lambda,\alpha,M)\widetilde{\mathcal{W}}(t)\mathcal{X}^{\sigma}(t,\lambda,\alpha,M)\Delta t \geq 0. \tag{4.32}$$

In addition, under Hypothesis 4.2, we have $\delta(\lambda)\operatorname{Im}(M) > 0$.

Proof. By identity (4.18), for any matrix $M \in D(\lambda)$, we have

$$\begin{aligned} 2\delta(\lambda)\operatorname{Im}(M) &= -\mathcal{E}(M) + 2|\operatorname{Im}(\lambda)| \int_a^b \mathcal{X}^{\sigma *}(t,\lambda,\alpha,M)\widetilde{\mathcal{W}}(t)\mathcal{X}^{\sigma}(t,\lambda,\alpha,M)\Delta t \\ &\geq 2|\operatorname{Im}(\lambda)| \int_a^b \mathcal{X}^{\sigma *}(t,\lambda,\alpha,M)\widetilde{\mathcal{W}}(t)\mathcal{X}^{\sigma}(t,\lambda,\alpha,M)\Delta t, \end{aligned} \tag{4.33}$$

which yields together with $\widetilde{\mathcal{W}}(t) \geq 0$ on $[a,\rho(b)]_{\mathbb{T}}$ the inequalities in (4.32). The last assertion in Theorem 4.13 is a simple consequence of Hypothesis 4.2. □

In the last part of this section we wish to study the effect of changing α, which is one of the parameters of the $M(\lambda)$-function and the Weyl solution $\mathcal{X}(\cdot,\lambda,\alpha,M)$, when α varies within the set Γ. For this purpose we will use the $M(\lambda)$-function with all its arguments in the following two statements.

Lemma 4.14. *Let $\alpha,\beta,\gamma \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Then*

$$M(\lambda,b,\alpha,\beta) = [\alpha\mathcal{J}\gamma^* + \alpha\gamma^* M(\lambda,b,\gamma,\beta)]\,[\alpha\gamma^* - \alpha\mathcal{J}\gamma^* M(\lambda,b,\gamma,\beta)]^{-1}. \tag{4.34}$$

Proof. Let $M(b,\lambda,\alpha,\beta)$ and $M(b,\lambda,\gamma,\beta)$ be given via (4.13), and consider the Weyl solutions $\mathcal{X}_\alpha(\cdot) := \mathcal{X}(\cdot,\lambda,\alpha,M(b,\lambda,\alpha,\beta))$ and $\mathcal{X}_\gamma(\cdot) := \mathcal{X}(\cdot,\lambda,\gamma,M(b,\lambda,\gamma,\beta))$ defined by (4.16) with $M = M(b,\lambda,\alpha,\beta)$ and $M = M(b,\lambda,\gamma,\beta)$, respectively. First we prove that the two Weyl solutions $\mathcal{X}_\alpha(\cdot)$ and $\mathcal{X}_\gamma(\cdot)$ differ by a constant nonsingular multiple. By definition, $\beta\mathcal{X}_\alpha(b) = 0$ and $\beta\mathcal{X}_\gamma(b) = 0$, which implies through (4.3) that $\mathcal{X}_\alpha(b) = \mathcal{J}\beta^* P_\alpha$ and $\mathcal{X}_\gamma(b) = \mathcal{J}\beta^* P_\gamma$

to Contents, page 1

for some matrices $P_\alpha, P_\gamma \in \mathbb{C}^{n\times n}$, which are invertible by Remark 4.11. This implies that $\mathcal{X}_\alpha(b)P_\alpha^{-1} = \mathcal{J}\beta^* = \mathcal{X}_\gamma(b)P_\gamma^{-1}$. Consequently, $\mathcal{X}_\alpha(b) = \mathcal{X}_\gamma(b)P$, where $P := P_\gamma^{-1}P_\alpha$. By the uniqueness of solutions of system $(\mathcal{S}_\lambda)$, see Theorem 3.4, we obtain that $\mathcal{X}_\alpha(\cdot) = \mathcal{X}_\gamma(\cdot)P$ on $[a,b]_{\mathbb{T}}$. Upon the evaluation at $t = a$ we get

$$\Psi(a,\lambda,\alpha)\begin{pmatrix} I \\ M(\lambda,b,\alpha,\beta)\end{pmatrix} = \Psi(a,\lambda,\gamma)\begin{pmatrix} I \\ M(\lambda,b,\gamma,\beta)\end{pmatrix}P. \tag{4.35}$$

Since the matrices $\Psi(a,\lambda,\alpha) = (\alpha^* \;\; -\mathcal{J}\alpha^*)$ and $\Psi(a,\lambda,\gamma) = (\gamma^* \;\; -\mathcal{J}\gamma^*)$ are unitary, it follows from (4.35) that

$$\begin{aligned}\begin{pmatrix} I \\ M(\lambda,b,\alpha,\beta)\end{pmatrix} &= \begin{pmatrix} \alpha \\ \alpha\mathcal{J}\end{pmatrix}(\gamma^* \;\; -\mathcal{J}\gamma^*)\begin{pmatrix} I \\ M(\lambda,b,\gamma,\beta)\end{pmatrix}P \\ &= \begin{pmatrix} \alpha\gamma^* - \alpha\mathcal{J}\gamma^* M(\lambda,b,\gamma,\beta) \\ \alpha\mathcal{J}\gamma^* + \alpha\gamma^* M(\lambda,b,\gamma,\beta)\end{pmatrix}P.\end{aligned} \tag{4.36}$$

The first row above yields that $P = [\alpha\gamma^* - \alpha\mathcal{J}\gamma^* M(\lambda,b,\gamma,\beta)]^{-1}$, while the second row is then written as identity (4.34). □

Corollary 4.15. *Let $\alpha,\beta,\gamma \in \Gamma$ and $\lambda \in \mathbb{C}\setminus\mathbb{R}$. With notation* (4.16) *and* (4.13) *we have*

$$\mathcal{X}(\cdot,\lambda,\alpha,M(\lambda,b,\alpha,\beta)) = \mathcal{X}(\cdot,\lambda,\gamma,M(\lambda,b,\gamma,\beta))\left[\alpha\gamma^* - \alpha\mathcal{J}\gamma^* M(\lambda,b,\gamma,\beta)\right]^{-1}. \tag{4.37}$$

Proof. The above identity follows from (4.35) and the formula for the matrix P from the end of the proof of Lemma 4.14. □

5. Geometric Properties of Weyl Disks

In this section we study the geometric properties of the Weyl disks as the point b moves through the interval $[a,\infty)_{\mathbb{T}}$. Our first result shows that the Weyl disks $D(\lambda,b)$ are nested. This statement generalizes the results in [11, Theorem 4.5], [66, Section 3.2.1], [9, equation (2.70)], [14, Theorem 3.1], [3, Theorem 4.4], and [5, Theorem 3.3(i)].

Theorem 5.1 (nesting property of Weyl disks). *Let $\alpha \in \Gamma$ and $\lambda \in \mathbb{C}\setminus\mathbb{R}$. Then*

$$D(\lambda,b_2) \subseteq D(\lambda,b_1), \quad \textit{for every } b_1,b_2 \in [a,\infty)_{\mathbb{T}},\; b_1 < b_2. \tag{5.1}$$

to Contents, page 1

Proof. Let $b_1, b_2 \in [a, \infty)_{\mathbb{T}}$ with $b_1 < b_2$, and take $M \in D(\lambda, b_2)$, that is, $\mathcal{E}(M, b_2) \leq 0$. From identity (4.18) with $b = b_1$ and later with $b = b_2$ and by using $\widetilde{\mathcal{W}}(\cdot) \geq 0$, we have

$$\begin{aligned}
\mathcal{E}(M, b_1) &\overset{(4.18)}{=} -2\delta(\lambda)\operatorname{Im}(M) + 2|\operatorname{Im}(\lambda)| \int_a^{b_1} \mathcal{X}^{\sigma *}(t, \lambda, \alpha, M)\widetilde{\mathcal{W}}(t)\mathcal{X}^{\sigma}(t, \lambda, \alpha, M)\Delta t \\
&\leq -2\delta(\lambda)\operatorname{Im}(M) + 2|\operatorname{Im}(\lambda)| \int_a^{b_2} \mathcal{X}^{\sigma *}(t, \lambda, \alpha, M)\widetilde{\mathcal{W}}(t)\mathcal{X}^{\sigma}(t, \lambda, \alpha, M)\Delta t \\
&\overset{(4.18)}{=} \mathcal{E}(M, b_2) \leq 0.
\end{aligned} \tag{5.2}$$

Therefore, by Definition 4.9, the matrix M belongs to $D(\lambda, b_1)$, which shows the result. □

Similarly for the regular case (Hypothesis 4.2) we now introduce the following assumption.

Hypothesis 5.2. There exists $b_0 \in (a, \infty)_{\mathbb{T}}$ such that Hypothesis 4.2 is satisfied with $b = b_0$; that is, inequality (4.8) holds with $b = b_0$ for every $\lambda \in \mathbb{C}$.

From Hypothesis 5.2 it follows by $\widetilde{\mathcal{W}}(\cdot) \geq 0$ that inequality (4.8) holds for every $b \in [b_0, \infty)_{\mathbb{T}}$.

For the study of the geometric properties of Weyl disks we will use the following representation:

$$\mathcal{E}(M, b) = i\delta(\lambda)\mathcal{X}^*(b, \lambda, \alpha, M)\mathcal{J}\mathcal{X}(b, \lambda, \alpha, M) = \begin{pmatrix} I & M^* \end{pmatrix} \begin{pmatrix} \mathcal{F}(b, \lambda, \alpha) & \mathcal{G}^*(b, \lambda, \alpha) \\ \mathcal{G}(b, \lambda, \alpha) & \mathcal{H}(b, \lambda, \alpha) \end{pmatrix} \begin{pmatrix} I \\ M \end{pmatrix}, \tag{5.3}$$

of the matrix $\mathcal{E}(M, b)$, where we define on $[a, \infty)_{\mathbb{T}}$ the $n \times n$ matrices

$$\begin{aligned}
\mathcal{F}(\cdot, \lambda, \alpha) &:= i\delta(\lambda)Z^*(\cdot, \lambda, \alpha)\mathcal{J}Z(\cdot, \lambda, \alpha), \\
\mathcal{G}(\cdot, \lambda, \alpha) &:= i\delta(\lambda)\widetilde{Z}^*(\cdot, \lambda, \alpha)\mathcal{J}Z(\cdot, \lambda, \alpha), \\
\mathcal{H}(\cdot, \lambda, \alpha) &:= i\delta(\lambda)\widetilde{Z}^*(\cdot, \lambda, \alpha)\mathcal{J}\widetilde{Z}(\cdot, \lambda, \alpha).
\end{aligned} \tag{5.4}$$

Since $\mathcal{E}(M, b)$ is Hermitian, it follows that $\mathcal{F}(\cdot, \lambda, \alpha)$ and $\mathcal{H}(\cdot, \lambda, \alpha)$ are also Hermitian. Moreover, by (4.7), we have $\mathcal{H}(a, \lambda, \alpha) = 0$. In addition, if $b \in [b_0, \infty)_{\mathbb{T}}$, then Corollary 3.7 and Hypothesis 5.2 yield for any $\lambda \in \mathbb{C} \setminus \mathbb{R}$

$$\mathcal{H}(b, \lambda, \alpha) = 2|\operatorname{Im}(\lambda)| \int_a^b \widetilde{Z}^{\sigma *}(t, \lambda, \alpha)\widetilde{\mathcal{W}}(t)\widetilde{Z}^{\sigma}(t, \lambda, \alpha)\Delta t > 0. \tag{5.5}$$

Therefore, $\mathcal{H}(b, \lambda, \alpha)$ is invertible (positive definite) for all $b \in [b_0, \infty)_{\mathbb{T}}$ and monotone nondecreasing as $b \to \infty$, with a consequence that $\mathcal{H}^{-1}(b, \lambda, \alpha)$ is monotone nonincreasing as $b \to \infty$. The following factorization of $\mathcal{E}(M, b)$ holds true; see also [2, equation (4.11)].

to Contents, page 1

Lemma 5.3. *Let $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. With the notation* (5.4), *for any $M \in \mathbb{C}^{n\times n}$ and $b \in [a,\infty)_{\mathbb{T}}$ we have*

$$\begin{aligned}\mathcal{E}(M,b) &= \mathcal{F}(b,\lambda,\alpha) - \mathcal{G}^*(b,\lambda,\alpha)\mathcal{H}^{-1}(b,\lambda,\alpha)\mathcal{G}(b,\lambda,\alpha) \\ &\quad + \left[\mathcal{G}^*(b,\lambda,\alpha)\mathcal{H}^{-1}(b,\lambda,\alpha) + M^*\right]\mathcal{H}(b,\lambda,\alpha)\left[\mathcal{H}^{-1}(b,\lambda,\alpha)\mathcal{G}(b,\lambda,\alpha) + M\right],\end{aligned} \tag{5.6}$$

whenever the matrix $\mathcal{H}(b,\lambda,\alpha)$ is invertible.

Proof. The result is shown by a direct calculation. □

The following identity is a generalization of its corresponding versions in [11, Lemma 4.3], [1, Lemma 3.3], [14, Proposition 3.2], [2, Lemma 4.2], [3, Lemma 4.6], and [5, Theorem 5.6].

Lemma 5.4. *Let $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. With the notation* (5.4), *for any $b \in [a,\infty)_{\mathbb{T}}$, we have*

$$\mathcal{G}^*(b,\lambda,\alpha)\mathcal{H}^{-1}(b,\lambda,\alpha)\mathcal{G}(b,\lambda,\alpha) - \mathcal{F}(b,\lambda,\alpha) = \mathcal{H}^{-1}\left(b,\overline{\lambda},\alpha\right), \tag{5.7}$$

whenever the matrices $\mathcal{H}(b,\lambda,\alpha)$ and $\mathcal{H}(b,\overline{\lambda},\alpha)$ are invertible.

Proof. In order to simplify and abbreviate the notation we introduce the matrices

$$\begin{aligned}&\mathcal{F} := \mathcal{F}(b,\lambda,\alpha), \qquad \mathcal{G} := \mathcal{G}(b,\lambda,\alpha), \qquad \mathcal{H} := \mathcal{H}(b,\lambda,\alpha), \\ &\widetilde{\mathcal{F}} := \mathcal{F}\left(b,\overline{\lambda},\alpha\right), \qquad \widetilde{\mathcal{G}} := \mathcal{G}\left(b,\overline{\lambda},\alpha\right), \qquad \widetilde{\mathcal{H}} := \mathcal{H}\left(b,\overline{\lambda},\alpha\right),\end{aligned} \tag{5.8}$$

and use the notation $Z(\lambda)$ and $\widetilde{Z}(\lambda)$ for $Z(b,\lambda,\alpha)$ and $\widetilde{Z}(b,\lambda,\alpha)$, respectively. Then, since $\mathcal{F}^* = \mathcal{F}$ and $\delta(\lambda)\delta(\overline{\lambda}) = -1$, we get the identities

$$\mathcal{G}^*\widetilde{\mathcal{F}} - \mathcal{F}^*\widetilde{\mathcal{G}} = Z^*(\lambda)\mathcal{J}\left[\widetilde{Z}(\lambda)Z^*\left(\overline{\lambda}\right) - Z(\lambda)\widetilde{Z}^*\left(\overline{\lambda}\right)\right]\mathcal{J}Z\left(\overline{\lambda}\right) \overset{(3.21)}{=} Z^*(\lambda)\mathcal{J}Z\left(\overline{\lambda}\right) \overset{(3.20)}{=} 0, \tag{5.9}$$

$$\mathcal{H}\widetilde{\mathcal{G}}^* - \mathcal{G}\mathcal{H}^* = \widetilde{Z}^*(\lambda)\mathcal{J}\left[\widetilde{Z}(\lambda)Z^*\left(\overline{\lambda}\right) - Z(\lambda)\widetilde{Z}^*\left(\overline{\lambda}\right)\right]\mathcal{J}\widetilde{Z}\left(\overline{\lambda}\right) \overset{(3.21)}{=} \widetilde{Z}^*(\lambda)\mathcal{J}\widetilde{Z}\left(\overline{\lambda}\right) \overset{(3.20)}{=} 0, \tag{5.10}$$

$$\mathcal{G}\widetilde{\mathcal{G}} - \mathcal{H}\widetilde{\mathcal{F}} = \widetilde{Z}^*(\lambda)\mathcal{J}\left[Z(\lambda)\widetilde{Z}^*\left(\overline{\lambda}\right) - \widetilde{Z}(\lambda)Z^*\left(\overline{\lambda}\right)\right]\mathcal{J}Z\left(\overline{\lambda}\right) \overset{(3.21)}{=} -\widetilde{Z}^*(\lambda)\mathcal{J}Z\left(\overline{\lambda}\right) \overset{(3.20)}{=} I, \tag{5.11}$$

$$\mathcal{G}^*\widetilde{\mathcal{G}}^* - \mathcal{F}\widetilde{\mathcal{H}} = Z^*(\lambda)\mathcal{J}\left[\widetilde{Z}(\lambda)Z^*\left(\overline{\lambda}\right) - Z(\lambda)\widetilde{Z}^*\left(\overline{\lambda}\right)\right]\mathcal{J}\widetilde{Z}\left(\overline{\lambda}\right) \overset{(3.21)}{=} Z^*(\lambda)\mathcal{J}\widetilde{Z}\left(\overline{\lambda}\right) \overset{(3.20)}{=} I. \tag{5.12}$$

Hence, by using that $\widetilde{\mathcal{H}}$ is Hermitian, we see that

$$\widetilde{\mathcal{H}}^{-1} \overset{(5.12)}{=} \mathcal{G}^*\widetilde{\mathcal{G}}^*\widetilde{\mathcal{H}}^{-1} - \mathcal{F} = \mathcal{G}^*\widetilde{\mathcal{G}}^*\widetilde{\mathcal{H}}^{*-1} - \mathcal{F} \overset{(5.10)}{=} \mathcal{G}^*\mathcal{H}^{-1}\mathcal{G} - \mathcal{F}. \tag{5.13}$$

Identity (5.7) is now proven. □

to Contents, page 1

Corollary 5.5. *Let $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Under Hypothesis 5.2, the matrix $\mathcal{H}(b,\lambda,\alpha)$ is invertible for every $b \in [b_0,\infty)_{\mathbb{T}}$, and for these values of b we have*

$$\mathcal{G}^*(b,\lambda,\alpha)\mathcal{H}^{-1}(b,\lambda,\alpha)\mathcal{G}(b,\lambda,\alpha) - \mathcal{F}(b,\lambda,\alpha) > 0. \tag{5.14}$$

Proof. Since $b \in [b_0,\infty)_{\mathbb{T}}$, then identity (5.5) yields that $\mathcal{H}(b,\lambda,\alpha) > 0$ and $\mathcal{H}(b,\overline{\lambda},\alpha) > 0$. Consequently, inequality (5.14) follows from (5.7) of Lemma 5.4. □

In the next result we justify the terminology for the sets $D(\lambda,b)$ and $C(\lambda,b)$ in Definition 4.9 to be called a "disk" and a "circle." It is a generalization of [14, Theorem 3.1], [2, Theorem 5.4], [5, Theorem 3.3(iii)]; see also [66, Theorem 3.5], [26, pages 70-71], [8, page 3485], [14, Proposition 3.3], [1, Theorem 3.3], [3, Theorem 4.8]. Consider the sets $\mathcal{V}$ and $\mathcal{U}$ of contractive and unitary matrices in $\mathbb{C}^{n\times n}$, respectively, that is,

$$\mathcal{V} := \{V \in \mathbb{C}^{n\times n},\ V^*V \le I\}, \qquad \mathcal{U} := \partial\mathcal{V} = \{U \in \mathbb{C}^{n\times n},\ U^*U = I\}. \tag{5.15}$$

The set $\mathcal{V}$ is known to be closed (in fact compact, since $\mathcal{V}$ is bounded) and convex.

Theorem 5.6. *Let $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Under Hypothesis 5.2, for every $b \in [b_0,\infty)_{\mathbb{T}}$, the Weyl disk and Weyl circle have the representations*

$$D(\lambda,b) = \left\{P(\lambda,b) + R(\lambda,b)VR\left(\overline{\lambda},b\right),\ V \in \mathcal{V}\right\}, \tag{5.16}$$

$$C(\lambda,b) = \left\{P(\lambda,b) + R(\lambda,b)UR\left(\overline{\lambda},b\right),\ U \in \mathcal{U}\right\}, \tag{5.17}$$

where, with the notation (5.4),

$$P(\lambda,b) := -\mathcal{H}^{-1}(\lambda,b,\alpha)\mathcal{G}(\lambda,b,\alpha), \qquad R(\lambda,b) := \mathcal{H}^{-1/2}(\lambda,b,\alpha). \tag{5.18}$$

Consequently, for every $b \in [b_0,\infty)_{\mathbb{T}}$, the sets $D(\lambda,b)$ are closed and convex.

The representations of $D(\lambda,b)$ and $C(\lambda,b)$ in (5.16) and (5.17) can be written as $D(\lambda,b) = P(\lambda,b) + R(\lambda,b)\mathcal{V}R(\overline{\lambda},b)$ and $C(\lambda,b) = P(\lambda,b) + R(\lambda,b)\mathcal{U}R(\overline{\lambda},b)$. The importance of the matrices $P(\lambda,b)$ and $R(\lambda,b)$ is justified in the following.

Definition 5.7. For $\alpha \in \Gamma$, $\lambda \in \mathbb{C} \setminus \mathbb{R}$, and $b \in [a,\infty)_{\mathbb{T}}$ such that $\mathcal{H}(\lambda,b,\alpha)$ and $\mathcal{H}(\overline{\lambda},b,\alpha)$ are positive definite, the matrix $P(\lambda,b)$ is called the center of the Weyl disk or the Weyl circle. The matrices $R(\lambda,b)$ and $R(\overline{\lambda},b)$ are called the *matrix radii* of the Weyl disk or the Weyl circle.

Proof of Theorem 5.6. By (5.5) and for any $b \in [b_0,\infty)_{\mathbb{T}}$, the matrices $\mathcal{H} := \mathcal{H}(\lambda,b,\alpha)$ and $\widetilde{\mathcal{H}} := \mathcal{H}(\overline{\lambda},b,\alpha)$ are positive definite, so that the matrices $P := P(\lambda,b)$, $R(\lambda) := R(\lambda,b)$, and

to Contents, page 1

$R(\overline{\lambda}) := R(\overline{\lambda}, b)$ are well defined. By Definition 4.9, for $M \in D(\lambda, b)$, we have $\mathcal{E}(M, b) \leq 0$, which in turn with notation (5.8) implies by Lemmas 5.3 and 5.4 that

$$\begin{aligned} &- R^2\left(\overline{\lambda}\right) + (M^* - P^*)R^{-2}(\lambda)(M - P) \\ &\overset{(5.7)}{=} \mathcal{F} - \mathcal{G}^*\mathcal{H}^{-1}\mathcal{G} + \left(\mathcal{H}^{-1}\mathcal{G} + M\right)^*\mathcal{H}\left(\mathcal{H}^{-1}\mathcal{G} + M\right) = \mathcal{E}(M, b) \leq 0. \end{aligned} \tag{5.19}$$

Therefore, the matrix

$$V := R^{-1}(\lambda)(M - P)R^{-1}\left(\overline{\lambda}\right), \tag{5.20}$$

satisfies $V^*V \leq I$. This relation between the matrices $M \in D(\lambda, b)$ and $V \in \mathcal{V}$ is bijective (more precisely, it is a homeomorphism), and the inverse to (5.20) is given by $M = P + R(\lambda)VR(\overline{\lambda})$. The latter formula proves that the Weyl disk $D(\lambda, b)$ has the representation in (5.16). Moreover, since by the definition $M \in C(\lambda, b)$ means that $\mathcal{E}(M, b) = 0$, it follows that the elements of the Weyl circle $C(\lambda, b)$ are in one-to-one correspondence with the matrices V defined in (5.20) which, similarly as in (5.19), now satisfy $V^*V = I$. Hence, the representation of $C(\lambda, b)$ in (5.17) follows. The fact that for $b \in [b_0, \infty)_{\mathbb{T}}$ the sets $D(\lambda, b)$ are closed and convex follows from the same properties of the set $\mathcal{V}$, being homeomorphic to $D(\lambda, b)$. □

6. Limiting Weyl Disk and Weyl Circle

In this section we study the limiting properties of the Weyl disk and Weyl circle and their center and matrix radii. Since under Hypothesis 5.2 the matrix function $\mathcal{H}(\cdot, \lambda, \alpha)$ is monotone nondecreasing as $b \to \infty$, it follows from the definition of $R(\lambda, b)$ and $R(\overline{\lambda}, b)$ in (5.18) that the two matrix functions $R(\lambda, \cdot)$ and $R(\overline{\lambda}, \cdot)$ are monotone nonincreasing for $b \to \infty$. Furthermore, since $R(\lambda, b)$ and $R(\overline{\lambda}, b)$ are Hermitian and positive definite for $b \in [b_0, \infty)_{\mathbb{T}}$, the limits

$$R_+(\lambda) := \lim_{b \to \infty} R(\lambda, b), \qquad R_+\left(\overline{\lambda}\right) := \lim_{b \to \infty} R\left(\overline{\lambda}, b\right), \tag{6.1}$$

exist and satisfy $R_+(\lambda) \geq 0$ and $R_+(\overline{\lambda}) \geq 0$. The index "+" in the above notation as well as in Definition 6.2 refers to the limiting disk at $+\infty$. In the following result we will see that the center $P(\lambda, b)$ also converges to a limiting matrix when $b \to \infty$. This is a generalization of [11, Theorem 4.7], [1, Theorem 3.5], [14, Proposition 3.5], [2, Theorem 4.5], and [3, Theorem 4.10].

Theorem 6.1. *Let $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Under Hypothesis 5.2, the center $P(\lambda, b)$ converges as $b \to \infty$ to a limiting matrix $P_+(\lambda) \in \mathbb{C}^{n\times n}$, that is,*

$$P_+(\lambda) := \lim_{b \to \infty} P(\lambda, b). \tag{6.2}$$

Proof. We prove that the matrix function $P(\lambda, \cdot)$ satisfies the Cauchy convergence criterion. Let $b_1, b_2 \in [b_0, \infty)_{\mathbb{T}}$ be given with $b_1 < b_2$. By Theorem 5.1, we have that $D(\lambda, b_2) \subseteq D(\lambda, b_1)$.

to Contents, page 1

Therefore, by (5.16) of Theorem 5.6, for a matrix $M \in D(\lambda, b_2)$, there are (unique) matrices $V_1, V_2 \in \mathcal{V}$ such that

$$M = P(\lambda, b_j) + R(\lambda, b_j)V_j R\left(\overline{\lambda}, b_j\right), \quad j \in \{1, 2\}. \tag{6.3}$$

Upon subtracting the two equations in (6.3), we get

$$P(\lambda, b_2) - P(\lambda, b_1) + R(\lambda, b_2)V_2 R\left(\overline{\lambda}, b_2\right) = R(\lambda, b_1)V_1 R\left(\overline{\lambda}, b_1\right). \tag{6.4}$$

This equation, when solved for V_1 in terms of V_2, has the form

$$V_1 = R^{-1}(\lambda, b_1)\left[P(\lambda, b_2) - P(\lambda, b_1) + R(\lambda, b_2)V_2 R\left(\overline{\lambda}, b_2\right)\right] R^{-1}\left(\overline{\lambda}, b_1\right) =: T(V_2), \tag{6.5}$$

which defines a continuous mapping $T : \mathcal{V} \to \mathcal{V}$, $T(V_2) = V_1$. Since $\mathcal{V}$ is compact, it follows that the mapping T has a fixed point in $\mathcal{V}$, that is, $T(V) = V$ for some matrix $V \in \mathcal{V}$. Equation $T(V) = V$ implies that

$$\begin{aligned} P(\lambda, b_2) - P(\lambda, b_1) &= R(\lambda, b_1)VR\left(\overline{\lambda}, b_1\right) - R(\lambda, b_2)VR\left(\overline{\lambda}, b_2\right) \\ &= [R(\lambda, b_1) - R(\lambda, b_2)]VR\left(\overline{\lambda}, b_1\right) - R(\lambda, b_2)V\left[R\left(\overline{\lambda}, b_1\right) - R\left(\overline{\lambda}, b_2\right)\right]. \end{aligned} \tag{6.6}$$

Hence, by $\|V\| \le 1$, we have

$$\|P(\lambda, b_2) - P(\lambda, b_1)\| \le \|R(\lambda, b_1) - R(\lambda, b_2)\| \left\|R\left(\overline{\lambda}, b_1\right)\right\| + \|R(\lambda, b_2)\| \left\|R\left(\overline{\lambda}, b_1\right) - R\left(\overline{\lambda}, b_2\right)\right\|. \tag{6.7}$$

Since the functions $R(\lambda, \cdot)$ and $R(\overline{\lambda}, \cdot)$ are monotone nonincreasing, they are bounded; that is, for some $K > 0$, we have $\|R(\lambda, b)\| \le K$ and $\|R(\overline{\lambda}, b)\| \le K$ for all $b \in [b_0, \infty)_{\mathbb{T}}$.

Let $\varepsilon > 0$ be arbitrary. The convergence of $R(\lambda, b)$ and $R(\overline{\lambda}, b)$ as $b \to \infty$ yields the existence of $b_3 \in [b_0, \infty)_{\mathbb{T}}$ such that for every $b_1, b_2 \in [b_3, \infty)_{\mathbb{T}}$ with $b_1 < b_2$ we have

$$\|R(\nu, b_1) - R(\nu, b_2)\| \le \frac{\varepsilon}{(2K)}, \quad \nu \in \left\{\lambda, \overline{\lambda}\right\}. \tag{6.8}$$

Using estimate (6.8) in inequality (6.7) we obtain for $b_2 > b_1 \ge b_3$

$$\|P(\lambda, b_2) - P(\lambda, b_1)\| < \frac{\varepsilon}{(2K)} \cdot K + \frac{\varepsilon}{(2K)} \cdot K = \varepsilon. \tag{6.9}$$

This means that the limit $P_+(\lambda) \in \mathbb{C}^{n\times n}$ in (6.2) exists, which completes the proof. $\square$

to Contents, page 1

By Theorems 5.1 and 5.6 we know that the Weyl disks $D(\lambda, b)$ are closed, convex, and nested as $b \to \infty$. Thereore the limit of $D(\lambda, b)$ as $b \to \infty$ is a closed, convex, and nonempty set. This motivates the following definition, which can be found in the special cases of system $(\mathcal{S}_\lambda)$ in [26, Theorem 3.3], [1, Theorem 3.6], [2, Definition 4.7], and [3, Theorem 4.12].

Definition 6.2 (limiting Weyl disk). Let $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Then the set

$$D_+(\lambda) := \bigcap_{b \in [a,\infty)_{\mathbb{T}}} D(\lambda, b), \tag{6.10}$$

is called the *limiting Weyl disk*. The matrix $P_+(\lambda)$ from Theorem 6.1 is called the *center* of $D_+(\lambda)$ and the matrices $R_+(\lambda)$ and $R_+(\overline{\lambda})$ from (6.1) its *matrix radii*.

As a consequence of Theorem 5.6, we obtain the following characterization of the limiting Weyl disk.

Corollary 6.3. *Let $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Under Hypothesis 5.2, we have*

$$D_+(\lambda) = P_+(\lambda) + R_+(\lambda) \mathcal{V} R_+\left(\overline{\lambda}\right), \tag{6.11}$$

where $\mathcal{V}$ is the set of all contractive matrices defined in (5.15).

From now on we assume that Hypothesis 5.2 holds, so that the limiting center $P_+(\lambda)$ and the limiting matrix radii $R_+(\lambda)$ and $R_+(\overline{\lambda})$ of $D_+(\lambda)$ are well defined.

Remark 6.4. By means of the nesting property of the disks (Theorem 5.1) and Theorems 4.10 and 4.12, it follows that the elements of the limiting Weyl disk $D_+(\lambda)$ are of the form

$$M_+(\lambda) \in D_+(\lambda), \quad M_+(\lambda) = \lim_{b \to \infty} M\big(\lambda, b, \alpha, \beta(b)\big), \tag{6.12}$$

where $\beta(b) \in \mathbb{C}^{n \times 2n}$ satisfies $\beta(b)\beta^*(b) = I$ and $i\delta(\lambda)\beta(b)\mathcal{J}\beta^*(b) \geq 0$ for all $b \in [a, \infty)$. Moreover, from Lemma 4.6, we conclude that

$$M_+^*(\lambda) = M_+\left(\overline{\lambda}\right). \tag{6.13}$$

A matrix $M_+(\lambda)$ from (6.12) is called a *half-line Weyl-Titchmarsh $M(\lambda)$-function*. Also, as noted in [2, Section 4], see also [8, Theorem 2.18], the function $M_+(\lambda)$ is a Herglotz function with rank n and has a certain integral representation (which will not be needed in this paper).

Our next result shows another characterization of the elements of $D_+(\lambda)$ in terms of the Weyl solution $\mathcal{X}(\cdot, \alpha, \lambda, M)$ defined in (4.16). This is a generalization of [11, page 671], [26, equation (3.2)], [1, Theorem 3.8(i)], [2, Theorem 4.8], and [3, Theorem 4.15].

to Contents, page 1

Theorem 6.5. *Let $\alpha \in \Gamma$, $\lambda \in \mathbb{C} \setminus \mathbb{R}$, and $M \in \mathbb{C}^{n\times n}$. The matrix M belongs to the limiting Weyl disk $D_+(\lambda)$ if and only if*

$$\int_a^\infty \mathcal{X}^{\sigma *}(t,\lambda,\alpha,M)\widetilde{\mathcal{W}}(t)\mathcal{X}^\sigma(t,\lambda,\alpha,M)\Delta t \le \frac{\operatorname{Im}(M)}{\operatorname{Im}(\lambda)}. \tag{6.14}$$

Proof. By Definition 6.2, we have $M \in D_+(\lambda)$ if and only if $M \in D(\lambda, b)$, that is, $\mathcal{E}(M,b) \le 0$, for all $b \in [a,\infty)_{\mathbb{T}}$. Therefore, by formula (4.18), we get

$$\int_a^b \mathcal{X}^{\sigma *}(t,\lambda,\alpha,M)\widetilde{\mathcal{W}}(t)\mathcal{X}^\sigma(t,\lambda,\alpha,M)\Delta t = \frac{\mathcal{E}(M,b)}{2|\operatorname{Im}(\lambda)|} + \frac{\delta(\lambda)\operatorname{Im}(M)}{|\operatorname{Im}(\lambda)|} \le \frac{\operatorname{Im}(M)}{\operatorname{Im}(\lambda)}, \tag{6.15}$$

for every $b \in [a,\infty)_{\mathbb{T}}$, which is equivalent to inequality (6.14). □

Remark 6.6. In [1, Definition 3.4], the notion of a boundary of the limiting Weyl disk $D_+(\lambda)$ is discussed. This would be a "limiting Weyl circle" according to Definitions 4.9 and 6.2. The description of matrices $M \in \mathbb{C}^{n\times n}$ laying on this boundary follows from Theorems 6.5 and 4.10, giving for such matrices M the equality

$$\int_a^\infty \mathcal{X}^{\sigma *}(t,\lambda,\alpha,M)\widetilde{\mathcal{W}}(t)\mathcal{X}^\sigma(t,\lambda,\alpha,M)\Delta t = \frac{\operatorname{Im}(M)}{\operatorname{Im}(\lambda)}. \tag{6.16}$$

Condition (6.16) is also equivalent to

$$\lim_{t\to\infty} \mathcal{X}^*(t,\lambda,\alpha,M)\mathcal{J}\mathcal{X}(t,\lambda,\alpha,M) = 0. \tag{6.17}$$

This is because, by (4.19) and the Lagrange identity (Corollary 3.6),

$$\begin{aligned}&\mathcal{X}^*(t,\lambda,\alpha,M)\mathcal{J}\mathcal{X}(t,\lambda,\alpha,M)\\ &\quad= 2i\operatorname{Im}(\lambda)\left[\frac{\operatorname{Im}(M)}{\operatorname{Im}(\lambda)} - \int_a^t \mathcal{X}^{\sigma *}(s,\lambda,\alpha,M)\widetilde{\mathcal{W}}(s)\mathcal{X}^\sigma(s,\lambda,\alpha,M)\Delta s\right],\end{aligned} \tag{6.18}$$

for every $t \in [a,\infty)_{\mathbb{T}}$. From this we can see that the integral on the right-hand side above converges for $t \to \infty$ and (6.16) holds if and only if condition (6.17) is satisfied. Characterizations (6.16) and (6.17) of the matrices M on the boundary of the limiting Weyl disk $D_+(\lambda)$ generalize the corresponding results in [1, Theorems 3.8(ii) and 3.9]; see also [14, Theorem 6.3].

Consider the linear space of square integrable C^1_{prd} functions

$$L^2_{\mathcal{W}} = L^2_{\mathcal{W}}[a,\infty)_{\mathbb{T}} := \left\{z : [a,\infty)_{\mathbb{T}} \longrightarrow \mathbb{C}^{2n},\ z \in C^1_{\mathrm{prd}},\ \|z(\cdot)\|_{\mathcal{W}} < \infty\right\}, \tag{6.19}$$

to Contents, page 1

where we define

$$\|z(\cdot)\|_{\mathcal{W}} := \sqrt{\langle z(\cdot), z(\cdot)\rangle_{\mathcal{W}}}, \qquad \langle z(\cdot), \tilde{z}(\cdot)\rangle_{\mathcal{W}} := \int_a^\infty z^{\sigma *}(t)\widetilde{\mathcal{W}}(t)\tilde{z}^{\sigma}(t)\Delta t. \tag{6.20}$$

In the following result we prove that the space $L^2_{\mathcal{W}}$ contains the columns of the Weyl solution $\mathcal{X}(\cdot, \lambda, \alpha, M)$ when M belongs to the limiting Weyl disk $D_+(\lambda)$. This implies that there are at least n linearly independent solutions of system $(\mathcal{S}_\lambda)$ in $L^2_{\mathcal{W}}$. This is a generalization of [11, Theorem 5.1], [14, Theorem 4.1], [2, Theorem 4.10], and [5, page 716].

Theorem 6.7. *Let $\alpha \in \Gamma$, $\lambda \in \mathbb{C} \setminus \mathbb{R}$, and $M \in D_+(\lambda)$. The columns of $\mathcal{X}(\cdot, \lambda, \alpha, M)$ form a linearly independent system of solutions of system $(\mathcal{S}_\lambda)$, each of which belongs to $L^2_{\mathcal{W}}$.*

Proof. Let $z_j(\cdot) := \mathcal{X}(\cdot, \lambda, \alpha, M)e_j$ for $j \in \{1, \dots, n\}$ be the columns of the Weyl solution $\mathcal{X}(\cdot, \lambda, \alpha, M)$, where e_j is the jth unit vector. We prove that the functions $z_1(\cdot), \dots, z_n(\cdot)$ are linearly independent. Assume that $\sum_{j=1}^n c_j z_j(\cdot) = 0$ on $[a, \infty)_{\mathbb{T}}$ for some $c_1, \dots, c_n \in \mathbb{C}$. Then $\mathcal{X}(\cdot, \lambda, \alpha, M)c = 0$, where $c := (c_1^*, \dots, c_n^*)^* \in \mathbb{C}^n$. It follows by (4.19) that

$$2ic^* \operatorname{Im}(M)c = c^* \mathcal{X}^*(a, \lambda, \alpha, M)\mathcal{J}\mathcal{X}(a, \lambda, \alpha, M)c = 0, \tag{6.21}$$

which implies the equality $c^*\delta(\lambda)\operatorname{Im}(M)c = 0$. Using that $M \in D_+(\lambda) \subseteq D(\lambda, b)$ for some $b \in [b_0, \infty)_{\mathbb{T}}$, we obtain from Theorem 4.13 that the matrix $\delta(\lambda)\operatorname{Im}(M)$ is positive definite. Hence, $c = 0$ so that the functions $z_1(\cdot), \dots, z_n(\cdot)$ are linearly independent. Finally, for every $j \in \{1, \dots, n\}$ we get from Theorem 6.5 the inequality

$$\|z_j(\cdot)\|^2_{\mathcal{W}} = \int_a^\infty z_j^{\sigma *}(t)\widetilde{\mathcal{W}}(t)z_j^{\sigma}(t)\Delta t \overset{(6.14)}{\leq} e_j^* \frac{\operatorname{Im}(M)}{\operatorname{Im}(\lambda)} e_j \leq \frac{\|\delta(\lambda)\operatorname{Im}(M)\|}{|\operatorname{Im}(\lambda)|} < \infty. \tag{6.22}$$

Thus, $z_j(\cdot) \in L^2_{\mathcal{W}}$ for every $j \in \{1, \dots, n\}$, and the proof is complete. □

Denote by $\mathcal{N}(\lambda)$ the linear space of all square integrable solutions of system $(\mathcal{S}_\lambda)$, that is,

$$\mathcal{N}(\lambda) := \left\{ z(\cdot) \in L^2_{\mathcal{W}},\ z(\cdot) \text{ solves } (\mathcal{S}_\lambda) \right\}. \tag{6.23}$$

Then as a consequence of Theorem 6.7 we obtain the estimate

$$\dim \mathcal{N}(\lambda) \geq n, \quad \text{for each } \lambda \in \mathbb{C} \setminus \mathbb{R}. \tag{6.24}$$

Next we discuss the situation when $\dim \mathcal{N}(\lambda) = n$ for some $\lambda \in \mathbb{C} \setminus \mathbb{R}$.

Lemma 6.8. *Let $\alpha \in \Gamma$, $\lambda \in \mathbb{C} \setminus \mathbb{R}$, and $\dim \mathcal{N}(\lambda) = n$. Then the matrix radii of the limiting Weyl disk $D_+(\lambda)$ satisfy $R_+(\lambda) = 0 = R_+(\overline{\lambda})$. Consequently, the set $D_+(\lambda)$ consists of the single matrix $M = P_+(\lambda)$, that is, the center of $D_+(\lambda)$, which is given by formula (6.2) of Theorem 6.1.*

to Contents, page 1

Proof. With the matrix radii $R_+(\lambda)$ and $R_+(\overline{\lambda})$ of $D_+(\lambda)$ defined in (6.1) and with the Weyl solution $\mathcal{X}(\cdot,\lambda,\alpha,M)$ given by a matrix $M \in D_+(\lambda)$, we observe that the columns of $\mathcal{X}(\cdot,\lambda,\alpha,M)$ form a basis of the space $\mathcal{N}(\lambda)$. Since the columns of the fundamental matrix $\Psi(\cdot,\lambda,\alpha) = (Z(\cdot,\lambda,\alpha)\ \ \widetilde{Z}(\cdot,\lambda,\alpha))$ span all solutions of system $(\mathcal{S}_\lambda)$, the definition of $\mathcal{X}(\cdot,\lambda,\alpha,M) = Z(\cdot,\lambda,\alpha) + \widetilde{Z}(\cdot,\lambda,\alpha)M$ yields that the columns of $\widetilde{Z}(\cdot,\lambda,\alpha)$ together with the columns of $\mathcal{X}(\cdot,\lambda,\alpha,M)$ form a basis of all solutions of system $(\mathcal{S}_\lambda)$. Hence, from $\dim \mathcal{N}(\lambda) = n$ and Theorem 6.7, we get that the columns of $\widetilde{Z}(\cdot,\lambda,\alpha)$ do not belong to $L^2_{\mathcal{W}}$. Consequently, by formula (5.5), the Hermitian matrix functions $\mathcal{H}(\cdot,\lambda,\alpha)$ and $\mathcal{H}(\cdot,\overline{\lambda},\alpha)$ defined in (5.4) are monotone nondecreasing on $[a,\infty)_{\mathbb{T}}$ without any upper bound; that is, their eigenvalues—being real—tend to ∞. Therefore, the functions $R(\lambda,\cdot)$ and $R(\overline{\lambda},\cdot)$ as defined in (5.18) have limits at ∞ equal to zero; that is, $R_+(\lambda) = 0$ and $R_+(\overline{\lambda}) = 0$. The fact that the set $D_+(\lambda) = \{P_+(\lambda)\}$ then follows from the characterization of $D_+(\lambda)$ in Corollary 6.3. $\square$

In the final result of this section, we establish another characterization of the matrices M from the limiting Weyl disk $D_+(\lambda)$. In comparison with Theorem 6.5, we now use a similar condition to the one in Theorem 4.12 for the regular spectral problem. However, a stronger assumption than Hypothesis 5.2 is now required for this result to hold; compare with [9, Lemma 2.21] and [2, Theorem 4.16].

Hypothesis 6.9. For every $a_0, b_0 \in (a,\infty)_{\mathbb{T}}$ with $a_0 < b_0$ and for every $\lambda \in \mathbb{C}$, we have

$$\int_{a_0}^{b_0} \Psi^{\sigma*}(t,\lambda,\alpha)\widetilde{\mathcal{W}}(t)\Psi^{\sigma}(t,\lambda,\alpha)\Delta t > 0. \tag{6.25}$$

Under Hypothesis 6.9, the Weyl disks $D(\lambda,b)$ converge to the limiting disk "monotonically" as $b \to \infty$; that is, the limiting Weyl disk $D_+(\lambda)$ is "open" in the sense that all of its elements lie inside $D_+(\lambda)$. This can be interpreted in view of Theorem 4.12 as $\mathcal{E}(M,t) < 0$ for all $t \in [a,\infty)_{\mathbb{T}}$.

Theorem 6.10. *Let $\alpha \in \Gamma$, $\lambda \in \mathbb{C}\setminus\mathbb{R}$, and $M \in \mathbb{C}^{n\times n}$. Under Hypothesis 6.9, the matrix $M \in D_+(\lambda)$ if and only if*

$$\mathcal{E}(M,t) < 0, \quad \forall t \in [a,\infty)_{\mathbb{T}}. \tag{6.26}$$

Proof. If condition (6.26) holds, then $M \in D_+(\lambda)$ follows from the definition of $D_+(\lambda)$. Conversely, suppose that $M \in D_+(\lambda)$, and let $t \in [a,\infty)_{\mathbb{T}}$ be given. Then for any $b \in (t,\infty)_{\mathbb{T}}$ we have by formula (4.18) that

$$\begin{aligned}\mathcal{E}(M,t) &= -2\delta(\lambda)\,\mathrm{Im}(M) + 2|\mathrm{Im}(\lambda)|\int_a^t \mathcal{X}^{\sigma*}(s,\lambda,\alpha,M)\widetilde{\mathcal{W}}(s)\mathcal{X}^{\sigma}(s,\lambda,\alpha,M)\Delta s \\ &= \mathcal{E}(M,b) - 2|\mathrm{Im}(\lambda)|\int_t^b \mathcal{X}^{\sigma*}(s,\lambda,\alpha,M)\widetilde{\mathcal{W}}(s)\mathcal{X}^{\sigma}(s,\lambda,\alpha,M)\Delta s,\end{aligned} \tag{6.27}$$

to Contents, page 1

where we used the property $\int_a^t f(s)\Delta s = \int_a^b f(s)\Delta s - \int_t^b f(s)\Delta s$. Since $M \in D_+(\lambda)$ is assumed, we have $M \in D(\lambda, b)$, that is, $\mathcal{E}(M,b) \leq 0$, while Hypothesis 6.9 implies the positivity of the integral over $[t,b]_{\mathbb{T}}$ in (6.27). Consequently, (6.27) yields that $\mathcal{E}(M,t) < 0$. $\square$

Remark 6.11. If we partition the Weyl solution $\mathcal{X}(\cdot,\lambda) := \mathcal{X}(\cdot,\lambda,\alpha,M)$ into two $n \times n$ blocks $\mathcal{X}_1(\cdot,\lambda)$ and $\mathcal{X}_2(\cdot,\lambda)$ as in (4.28), then condition (6.26) can be written as

$$\delta(\lambda)\,\mathrm{Im}\big(\mathcal{X}_1^*(t,\lambda)\mathcal{X}_2(t,\lambda)\big) > 0, \quad \forall t \in [a,\infty)_{\mathbb{T}}. \tag{6.28}$$

Therefore, by Remark 2.2, the matrices $\mathcal{X}_1(t,\lambda)$ and $\mathcal{X}_2(t,\lambda)$ are invertible for all $t \in [a,\infty)_{\mathbb{T}}$. A standard argument then yields that the quotient $Q(\cdot,\lambda) := \mathcal{X}_2(\cdot,\lambda)\mathcal{X}_1^{-1}(\cdot,\lambda)$ satisfies the *Riccati matrix equation* (suppressing the argument t in the coefficients)

$$Q^{\Delta} - (\mathcal{C} + \mathcal{D}Q) + Q^{\sigma}(\mathcal{A} + \mathcal{B}Q) + \lambda\mathcal{W}\big[I + \mu(\mathcal{A} + \mathcal{B}Q)\big] = 0, \quad t \in [a,\infty)_{\mathbb{T}}, \tag{6.29}$$

see [57, Theorem 3], [48, Section 6], and [49].

7. Limit Point and Limit Circle Criteria

Throughout this section we assume that Hypothesis 5.2 is satisfied. The results from Theorem 6.7 and Lemma 6.8 motivate the following terminology; compare with [4, page 75], [43, Definition 1.2] in the time scales scalar case $n = 1$, with [8, page 3486], [36, page 1668], [30, page 274], [38, Definition 3.1], [37, Definition 1], [67, page 2826] in the continuous case, and with [14, Definition 5.1], [2, Definition 4.12] in the discrete case.

Definition 7.1 (limit point and limit circle case for system $(\mathcal{S}_\lambda)$). The system $(\mathcal{S}_\lambda)$ is said to be in the *limit point case* at ∞ (or of the *limit point type*) if

$$\dim \mathcal{N}(\lambda) = n, \quad \forall \lambda \in \mathbb{C} \setminus \mathbb{R}. \tag{7.1}$$

The system $(\mathcal{S}_\lambda)$ is said to be in the *limit circle case* at ∞ (or of the *limit circle type*) if

$$\dim \mathcal{N}(\lambda) = 2n, \quad \forall \lambda \in \mathbb{C} \setminus \mathbb{R}. \tag{7.2}$$

Remark 7.2. According to Remark 6.4 (in which $\beta(b) \equiv \beta$), the center $P_+(\lambda)$ of the limiting Weyl disk $D_+(\lambda)$ can be expressed in the limit point case as

$$P_+(\lambda) = M_+(\lambda) = \lim_{b\to\infty} M\big(\lambda, b, \alpha, \beta\big), \tag{7.3}$$

where $\beta \in \Gamma$ is arbitrary but fixed.

Next we establish the first main result of this section. Its continuous time version can be found in [30, Theorem 2.1], [11, Theorem 8.5] and the discrete time version in [9, Lemma 3.2], [2, Theorem 4.13].

to Contents, page 1

Theorem 7.3. *Let the system* $(\mathcal{S}_\lambda)$ *be in the limit point or limit circle case, fix* $\alpha \in \Gamma$, *and let* $\lambda, \nu \in \mathbb{C} \setminus \mathbb{R}$. *Then*

$$\lim_{t\to\infty} \mathcal{X}_+^*(t,\lambda,\alpha,M_+(\lambda))\mathcal{J}\mathcal{X}_+(t,\nu,\alpha,M_+(\nu)) = 0, \tag{7.4}$$

where $\mathcal{X}_+(\cdot,\lambda,\alpha,M_+(\lambda))$ *and* $\mathcal{X}_+(\cdot,\nu,\alpha,M_+(\nu))$ *are the Weyl solutions of* $(\mathcal{S}_\lambda)$ *and* $(\mathcal{S}_\nu)$, *respectively, defined by* (4.16) *through the matrices* $M_+(\lambda)$ *and* $M_+(\nu)$, *which are determined by the limit in* (6.12).

Proof. For every $t \in [a,\infty)_\mathbb{T}$ and matrices $\beta(t) \in \mathbb{C}^{n\times 2n}$ such that $\beta(t)\beta^*(t) = I$ and $i\delta(\lambda)\beta(t)\mathcal{J}\beta^*(t) \geq 0$ and for $\kappa \in \{\lambda,\nu\}$, we define the matrix (compare with Definition 4.5)

$$M(\kappa,t,\alpha,\beta(t)) := -\left[\beta(t)\widetilde{Z}(t,\kappa,\alpha)\right]^{-1}\beta(t)Z(t,\kappa,\alpha). \tag{7.5}$$

Then, by Theorems 4.10 and 4.12, we have $M(\kappa,t,\alpha,\beta(t)) \in D(\kappa,t)$. Following the notation in (4.16), we consider the Weyl solutions $\mathcal{X}(\cdot,\kappa) := \mathcal{X}(\cdot,\kappa,\alpha,M(\kappa,t,\alpha,\beta(\cdot)))$. Similarly, let $\mathcal{X}_+(\cdot,\kappa) := \mathcal{X}(\cdot,\kappa,\alpha,M_+(\kappa))$ be the Weyl solutions corresponding to the matrices $M_+(\kappa) \in D_+(\kappa)$ from the statement of this theorem.

First assume that the system $(\mathcal{S}_\lambda)$ is of the limit point type. In this case, by Remark 7.2, we may take $\beta(t) \in \Gamma$ for all $t \in [a,\infty)_\mathbb{T}$. Hence, from Theorem 4.10, we get that $\beta(\cdot)\mathcal{X}(\cdot,\kappa) = 0$ on $[a,\infty)_\mathbb{T}$. By (4.3), for each $t \in [a,\infty)_\mathbb{T}$ and $\kappa \in \{\lambda,\nu\}$, there is a matrix $Q_\kappa(t) \in \mathbb{C}^{n\times n}$ such that $\mathcal{X}(\cdot,\kappa) = \mathcal{J}\beta^*(\cdot)Q_\kappa(\cdot)$ on $[a,\infty)_\mathbb{T}$. Hence, we have on $[a,\infty)_\mathbb{T}$

$$\begin{aligned}
&\mathcal{X}_+^*(t,\lambda)\mathcal{J}\mathcal{X}_+(t,\nu) + F(t,\lambda,\nu,\beta(t)) + G(t,\lambda,\nu,\beta(t)) \\
&\quad = \mathcal{X}^*(t,\lambda)\mathcal{J}\mathcal{X}(t,\nu) = Q_\lambda^*(t)\beta(t)\mathcal{J}\beta^*(t)Q_\nu(t) = 0,
\end{aligned} \tag{7.6}$$

where we define

$$\begin{aligned}
F(t,\lambda,\nu,\beta(t)) &:= \mathcal{X}_+^*(t,\lambda)\mathcal{J}\widetilde{Z}(t,\nu,\alpha)\left[M(\nu,t,\alpha,\beta(t)) - M_+(\nu)\right], \\
G(t,\lambda,\nu,\beta(t)) &:= \left[M^*(\lambda,t,\alpha,\beta(t)) - M_+^*(\lambda)\right]\widetilde{Z}^*(t,\lambda,\alpha)\mathcal{J}\mathcal{X}(t,\nu).
\end{aligned} \tag{7.7}$$

If we show that

$$\lim_{t\to\infty} F(t,\lambda,\nu,\beta(t)) = 0, \qquad \lim_{t\to\infty} G(t,\lambda,\nu,\beta(t)) = 0, \tag{7.8}$$

then (7.6) implies the result claimed in (7.4). First we prove the second limit in (7.8). Pick any $t \in [b_0,\infty)_\mathbb{T}$. By Theorem 5.6, Corollary 6.3, and $D_+(\lambda) \subseteq D(\lambda,t)$, we have

$$M(\lambda,t,\alpha,\beta(t)) = P(\lambda,t) + R(\lambda,t)U(t)R\left(\overline{\lambda},t\right), \qquad M_+(\lambda) = P(\lambda,t) + R(\lambda,t)V(t)R\left(\overline{\lambda},t\right), \tag{7.9}$$

to Contents, page 1

where $U(t) \in \mathcal{U}$ and $V(t) \in \mathcal{V}$. Therefore,

$$M(\lambda, t, \alpha, \beta(t)) - M_+(\lambda) = R(\lambda, t)[U(t) - V(t)]R\left(\overline{\lambda}, t\right). \tag{7.10}$$

Since $\widetilde{Z}(\cdot, \lambda, \alpha)$ and $\mathcal{X}(\cdot, \nu)$ are, respectively, solutions of systems $(\mathcal{S}_\lambda)$ and $(\mathcal{S}_\nu)$ which satisfy $\widetilde{Z}^*(a, \lambda, \alpha)\mathcal{J}\mathcal{X}(a, \nu) = -I$, it follows from Corollary 3.6 that

$$\widetilde{Z}^*(t, \lambda, \alpha)\mathcal{J}\mathcal{X}(t, \nu) = -I + \left(\overline{\lambda} - \nu\right) \int_a^t \widetilde{Z}^{\sigma *}(s, \lambda, \alpha)\widetilde{\mathcal{W}}(s)\mathcal{X}^\sigma(s, \nu)\Delta s. \tag{7.11}$$

Hence, we can write

$$G(t, \lambda, \nu, \beta(t)) = R\left(\overline{\lambda}, t\right)[U^*(t) - V^*(t)]R(\lambda, t)\left[\left(\overline{\lambda} - \nu\right) \int_a^t \widetilde{Z}^{\sigma *}(s, \lambda, \alpha)\widetilde{\mathcal{W}}(s)\mathcal{X}^\sigma(s, \nu)\Delta s - I\right], \tag{7.12}$$

where we used the Hermitian property of $R(\lambda, t)$ and $R(\overline{\lambda}, t)$. Since we now assume that system $(\mathcal{S}_\lambda)$ is in the limit point case, we know from Lemma 6.8 that $\lim_{t \to \infty} R(\lambda, t) = 0$ and $\lim_{t \to \infty} R(\overline{\lambda}, t) = 0$. Therefore, in order to establish (7.8)(ii), it is sufficient to show that

$$R(\lambda, t) \int_a^t \widetilde{Z}^{\sigma *}(s, \lambda, \alpha)\widetilde{\mathcal{W}}(s)\mathcal{X}^\sigma(s, \nu)\Delta s, \tag{7.13}$$

is bounded for $t \in [b_0, \infty)_{\mathbb{T}}$. Let $\eta \in \mathbb{C}^n$ be a unit vector, and denote by $\mathcal{X}_j(\cdot, \nu) := \mathcal{X}(\cdot, \nu)e_j$ the jth column of $\mathcal{X}(\cdot, \nu)$ for $j \in \{1, \dots, n\}$. With the definition of $R(\lambda, \cdot)$ in (5.18) we have

$$\begin{aligned}
&\left|\int_a^t \eta^* R(\lambda, s)\widetilde{Z}^{\sigma *}(s, \lambda, \alpha)\widetilde{\mathcal{W}}(s)\mathcal{X}_j^\sigma(s, \nu)\Delta s\right| \\
&\quad \le \int_a^t \left|\widetilde{\mathcal{W}}^{1/2}(s)\widetilde{Z}^{\sigma *}(s, \lambda, \alpha)R(\lambda, s)\eta\right|\left|\widetilde{\mathcal{W}}^{1/2}(s)\mathcal{X}_j^\sigma(s, \nu)\right|\Delta s \\
&\quad \overset{\text{C-S}}{\le} \left(\int_a^t \eta^* R(\lambda, s)\widetilde{Z}^{\sigma *}(s, \lambda, \alpha)\widetilde{\mathcal{W}}(s)\widetilde{Z}^\sigma(s, \lambda, \alpha)R(\lambda, s)\eta\Delta s\right)^{1/2} \\
&\qquad \times \left(\int_a^t \mathcal{X}_j^{\sigma *}(s, \nu)\widetilde{\mathcal{W}}(s)\mathcal{X}_j^\sigma(s, \nu)\Delta s\right)^{1/2},
\end{aligned} \tag{7.14}$$

where the last step follows from the Cauchy-Schwarz inequality (C-S) on time scales. From (5.5) we obtain

$$\mathcal{H}^{-1/2}(t, \lambda, \alpha) \int_a^t \widetilde{Z}^{\sigma *}(s, \lambda, \alpha)\widetilde{\mathcal{W}}(s)\widetilde{Z}^\sigma(s, \lambda, \alpha)\Delta s \mathcal{H}^{-1/2}(t, \lambda, \alpha) = \frac{1}{2|\operatorname{Im}(\lambda)|}I, \tag{7.15}$$

to Contents, page 1

so that the first term in the product in (7.14) is bounded by $1/\sqrt{2|\operatorname{Im}(\lambda)|}$. Moreover, from formula (4.18) we get that the second term in the product in (7.14) is bounded by the number $[e_j^* \operatorname{Im}(M(\nu,t,\alpha,\beta(t)))e_j]/\operatorname{Im}(\nu)$. Hence, upon recalling the limit in (6.12), we conclude that the product in (7.14) is bounded by

$$\frac{1}{2|\operatorname{Im}(\lambda)|}\cdot\frac{e_j^* \operatorname{Im}(M_+(\nu))e_j}{\operatorname{Im}(\nu)}, \tag{7.16}$$

which is independent of t. Consequently, the second limit in (7.8) is established. The first limit in (7.8) is then proven in a similar manner. The proof for the limit point case is finished.

If the system $(\mathcal{S}_\lambda)$ is in the limit circle case, then for $\kappa \in \{\lambda,\nu\}$ the columns of $\widetilde{Z}(\cdot,\kappa,\alpha)$ and $\mathcal{X}_+(\cdot,\kappa)$ belong to $L^2_{\mathcal{W}}$; hence, they are bounded in the $L^2_{\mathcal{W}}$ norm. In this case the limits in (7.8) easily follow from the limit (6.12) for $M_+(\kappa)$, $\kappa \in \{\lambda,\nu\}$. $\square$

In the next result we provide a characterization of the system $(\mathcal{S}_\lambda)$ being of the limit point type. Special cases of this statement can be found, for example, in [14, Theorem 6.12] and [2, Theorem 4.14].

Theorem 7.4. *Let $\alpha \in \Gamma$. The system $(\mathcal{S}_\lambda)$ is in the limit point case if and only if, for every $\lambda \in \mathbb{C}\setminus\mathbb{R}$ and every square integrable solutions $z_1(\cdot,\lambda)$ and $z_2(\cdot,\bar{\lambda})$ of $(\mathcal{S}_\lambda)$ and $(\mathcal{S}_{\bar{\lambda}})$, respectively, we have*

$$z_1^*(t,\lambda)\mathcal{J}z_2\left(t,\bar{\lambda}\right) = 0, \quad \forall t \in [b_0,\infty)_{\mathbb{T}}. \tag{7.17}$$

Proof. Let $(\mathcal{S}_\lambda)$ be in the limit point case. Fix any $\lambda \in \mathbb{C}\setminus\mathbb{R}$, and suppose that $z_1(\cdot,\lambda)$ and $z_2(\cdot,\bar{\lambda})$ are solutions of $(\mathcal{S}_\lambda)$ and $(\mathcal{S}_{\bar{\lambda}})$, respectively. Then, by Theorem 6.7 and Remark 6.4, there are vectors $\xi_1,\xi_2 \in \mathbb{C}^n$ such that $z_1(\cdot,\lambda) = \mathcal{X}_+(\cdot,\lambda)\xi_1$ and $z_2(\cdot,\bar{\lambda}) = \mathcal{X}_+(\cdot,\bar{\lambda})\xi_2$ on $[a,\infty)_{\mathbb{T}}$, where $\mathcal{X}_+(\cdot,\kappa) := \mathcal{X}_+(\cdot,\kappa,\alpha,M_+(\kappa))$ are the Weyl solutions corresponding to some matrices $M_+(\kappa) \in D_+(\kappa)$ for $\kappa \in \{\lambda,\bar{\lambda}\}$. In fact, by Lemma 6.8, the matrix $M_+(\kappa)$ is equal to the center of the disk $D_+(\kappa)$. It follows that for any $t \in [b_0,\infty)_{\mathbb{T}}$ equality

$$\begin{aligned}
&\mathcal{X}_+^*(t,\lambda)\mathcal{J}\mathcal{X}_+\left(t,\bar{\lambda}\right) \\
&\quad\overset{(4.16)}{=} (I \;\; M_+^*(\lambda))\Psi^*(t,\lambda,\alpha)\mathcal{J}\Psi\left(t,\bar{\lambda},\alpha\right)\left(I \;\; M_+^*(\bar{\lambda})\right)^* \overset{(3.19)(\mathrm{i})}{=} M_+^*\left(\bar{\lambda}\right) - M_+^*(\lambda) \overset{(6.13)}{=} 0,
\end{aligned} \tag{7.18}$$

holds, so that (7.17) is established. Conversely, let $\nu \in \mathbb{C}\setminus\mathbb{R}$ be arbitrary but fixed, set $\lambda := \bar{\nu}$, and suppose that, for every square integrable solutions $z_1(\cdot,\lambda)$ and $z_2(\cdot,\nu)$ of $(\mathcal{S}_\lambda)$ and $(\mathcal{S}_\nu)$, condition (7.17) is satisfied. From Theorem 6.7 we know that for $M_+(\kappa) \in D_+(\kappa)$ the columns $\mathcal{X}_+^{[j]}(\cdot,\kappa)$, $j \in \{1,\dots,n\}$, of the Weyl solution $\mathcal{X}_+(\cdot,\kappa)$ are linearly independent square integrable solutions of $(\mathcal{S}_\kappa)$, $\kappa \in \{\lambda,\nu\}$. Therefore, $\dim\mathcal{N}(\lambda) \ge n$, and $\dim\mathcal{N}(\nu) \ge n$. Moreover, by identity (3.19)(i), we have

$$\mathcal{X}_+^*(t,\lambda)\mathcal{J}\mathcal{X}_+^{[j]}(t,\nu) = 0, \quad \forall t \in [b_0,\infty)_{\mathbb{T}},\ j \in \{1,\dots,n\}. \tag{7.19}$$

to Contents, page 1

Let $z(\cdot,\nu)$ be any square integrable solution of system $(\mathcal{S}_\nu)$. Then, by our assumption (7.17),

$$\mathcal{X}_+^*(t,\lambda)\mathcal{J}z(t,\nu)=0, \quad \forall t\in[b_0,\infty)_{\mathbb{T}}. \tag{7.20}$$

From (7.19) and (7.20) it follows that the vectors $\mathcal{X}_+^{[j]}(a,\nu)$, $j\in\{1,\dots,n\}$, and $z(a,\nu)$ are solutions of the linear homogeneous system

$$\mathcal{X}_+^*(a,\lambda)\mathcal{J}\eta=0. \tag{7.21}$$

Since, by Theorem 6.7, the vectors $\mathcal{X}_+^{[j]}(a,\nu)$ for $j\in\{1,\dots,n\}$ represent a basis of the solution space of system (7.21), there exists a vector $\xi\in\mathbb{C}^n$ such that $z(a,\nu)=\mathcal{X}_+(a,\nu)\xi$. By the uniqueness of solutions of system $(\mathcal{S}_\nu)$ we then get $z(\cdot,\nu)=\mathcal{X}_+(\cdot,\nu)\xi$ on $[a,\infty)_{\mathbb{T}}$. Hence, the solution $z(\cdot,\nu)$ is square integrable and $\dim\mathcal{N}(\nu)=n$. Since $\nu\in\mathbb{C}\setminus\mathbb{R}$ was arbitrary, it follows that the system $(\mathcal{S}_\lambda)$ is in the limit point case. □

As a consequence of the above result, we obtain a characterization of the limit point case in terms of a condition similar to (7.17), but using a limit. This statement is a generalization of [30, Corollary 2.3], [9, Corollary 3.3], [14, Theorem 6.14], [2, Corollary 4.15], [1, Theorem 3.9], [3, Theorem 4.16].

Corollary 7.5. *Let $\alpha\in\Gamma$. The system $(\mathcal{S}_\lambda)$ is in the limit point case if and only if, for every $\lambda,\nu\in\mathbb{C}\setminus\mathbb{R}$ and every square integrable solutions $z_1(\cdot,\lambda)$ and $z_2(\cdot,\nu)$ of $(\mathcal{S}_\lambda)$ and $(\mathcal{S}_\nu)$, respectively, we have*

$$\lim_{t\to\infty} z_1^*(t,\lambda)\mathcal{J}z_2(t,\nu)=0. \tag{7.22}$$

Proof. The necessity follows directly from Theorem 7.3. Conversely, assume that condition (7.22) holds for every $\lambda,\nu\in\mathbb{C}\setminus\mathbb{R}$ and every square integrable solutions $z_1(\cdot,\lambda)$ and $z_2(\cdot,\nu)$ of $(\mathcal{S}_\lambda)$ and $(\mathcal{S}_\nu)$. Fix $\lambda\in\mathbb{C}\setminus\mathbb{R}$, and set $\nu:=\overline{\lambda}$. By Corollary 3.7 we know that $z_1^*(\cdot,\lambda)\mathcal{J}z_2(\cdot,\nu)$ is constant on $[a,\infty)_{\mathbb{T}}$. Therefore, by using condition (7.22), we can see that identity (7.17) must be satisfied, which yields by Theorem 7.4 that the system $(\mathcal{S}_\lambda)$ is of the limit point type. □

8. Nonhomogeneous Time Scale Symplectic Systems

In this section we consider the nonhomogeneous time scale symplectic system

$$z^\Delta(t,\lambda)=\mathcal{S}(t,\lambda)z(t,\lambda)-\mathcal{J}\widetilde{\mathcal{W}}(t)f^\sigma(t), \quad t\in[a,\infty)_{\mathbb{T}}, \tag{8.1}$$

where the matrix function $\mathcal{S}(\cdot,\lambda)$ and $\widetilde{\mathcal{W}}(\cdot)$ are defined in (3.3) and (3.1), $f\in L^2_{\mathcal{W}}$, and where the associated homogeneous system $(\mathcal{S}_\lambda)$ is either of the limit point or limit circle type at ∞. Together with system (8.1) we consider a second system of the same form but with a different spectral parameter and a different nonhomogeneous term

$$y^\Delta(t,\nu)=\mathcal{S}(t,\nu)y(t,\nu)-\mathcal{J}\widetilde{\mathcal{W}}(t)g^\sigma(t), \quad t\in[a,\infty)_{\mathbb{T}}, \tag{8.2}$$

with $g\in L^2_{\mathcal{W}}$. The following is a generalization of Theorem 3.5 to nonhomogeneous systems.

to Contents, page 1

Theorem 8.1 (Lagrange identity). *Let $\lambda, \nu \in \mathbb{C}$ and $m \in \mathbb{N}$ be given. If $z(\cdot,\lambda)$ and $y(\cdot,\nu)$ are $2n \times m$ solutions of systems* (8.1) *and* (8.2), *respectively, then*

$$\begin{aligned}&\left[z^*(t,\lambda)\mathcal{J}y(t,\nu)\right]^{\Delta}\\&\quad=\left(\overline{\lambda}-\nu\right)z^{\sigma*}(t,\lambda)\widetilde{\mathcal{W}}(t)y^{\sigma}(t,\nu)-f^{\sigma*}(t)\widetilde{\mathcal{W}}(t)y^{\sigma}(t,\nu)+z^{\sigma*}(t,\lambda)\widetilde{\mathcal{W}}(t)g^{\sigma}(t),\quad t\in[a,\infty)_{\mathbb{T}}.\end{aligned} \tag{8.3}$$

Proof. Formula (8.3) follows by the product rule (2.1) with the aid of the relation

$$z^{\sigma}(t,\lambda)=\left[I+\mu(t)\mathcal{S}(t,\lambda)\right]z(t,\lambda)+\mu(t)\widetilde{\mathcal{W}}(t)f^{\sigma}(t), \tag{8.4}$$

and identity (3.6). □

For $\alpha\in\Gamma$, $\lambda\in\mathbb{C}\setminus\mathbb{R}$, and $t,s\in[a,\infty)_{\mathbb{T}}$, we define the function

$$G(t,s,\lambda,\alpha):=\begin{cases}\widetilde{Z}(t,\lambda,\alpha)\mathcal{X}_+^*\left(s,\overline{\lambda},\alpha\right), & \text{for } t\in[a,s)_{\mathbb{T}},\\ \mathcal{X}_+(t,\lambda,\alpha)\widetilde{Z}^*\left(s,\overline{\lambda},\alpha\right), & \text{for } t\in[s,\infty)_{\mathbb{T}},\end{cases} \tag{8.5}$$

where $\widetilde{Z}(\cdot,\lambda,\alpha)$ is the solution of system $(\mathcal{S}_\lambda)$ given in (4.10), that is, $\widetilde{Z}(a,\lambda,\alpha)=-\mathcal{J}\alpha^*$, and $\mathcal{X}_+(\cdot,\lambda,\alpha):=\mathcal{X}(\cdot,\lambda,\alpha,M_+(\lambda))$ is the Weyl solution of $(\mathcal{S}_\lambda)$ as in (4.16) determined by a matrix $M_+(\lambda)\in D_+(\lambda)$. This matrix $M_+(\lambda)\in D_+(\lambda)$ is arbitrary but fixed throughout this section. By interchanging the order of the arguments t and s, we have

$$G(t,s,\lambda,\alpha)=\begin{cases}\mathcal{X}_+(t,\lambda,\alpha)\widetilde{Z}^*\left(s,\overline{\lambda},\alpha\right), & \text{for } s\in[a,t]_{\mathbb{T}},\\ \widetilde{Z}(t,\lambda,\alpha)\mathcal{X}_+^*\left(s,\overline{\lambda},\alpha\right), & \text{for } s\in(t,\infty)_{\mathbb{T}}.\end{cases} \tag{8.6}$$

In the literature the function $G(\cdot,\cdot,\lambda,\alpha)$ is called a resolvent kernel, compare with [30, page 283], [32, page 15], [2, equation (5.4)], and in this section it will play a role of the Green function.

Lemma 8.2. *Let $\alpha\in\Gamma$ and $\lambda\in\mathbb{C}\setminus\mathbb{R}$. Then*

$$\mathcal{X}_+(t,\lambda,\alpha)\widetilde{Z}^*\left(t,\overline{\lambda},\alpha\right)-\widetilde{Z}(t,\lambda,\alpha)\mathcal{X}_+^*\left(t,\overline{\lambda},\alpha\right)=\mathcal{J},\quad \forall t\in[a,\infty)_{\mathbb{T}}. \tag{8.7}$$

Proof. Identity (8.7) follows by a direct calculation from the definition of $\mathcal{X}_+(\cdot,\lambda,\alpha)$ via (4.16) with a matrix $M_+(\lambda)\in D_+(\lambda)$ by using formulas (3.21) and (6.13). □

In the next lemma we summarize the properties of the function $G(\cdot,\cdot,\lambda,\alpha)$, which together with Proposition 8.4 and Theorem 8.5 justifies the terminology "Green function" of the system (8.1); compare with [68, Section 4]. A discrete version of the following result can be found in [2, Lemma 5.1].

to Contents, page 1

Lemma 8.3. *Let $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. The function $G(\cdot,\cdot,\lambda,\alpha)$ has the following properties:*

(i) $G^*(t,s,\lambda,\alpha) = G(s,t,\overline{\lambda},\alpha)$ *for every* $t,s \in [a,\infty)_{\mathbb{T}}$, $t \neq s$,

(ii) $G^*(t,t,\lambda,\alpha) = G(t,t,\overline{\lambda},\alpha) - \mathcal{J}$ *for every* $t \in [a,\infty)_{\mathbb{T}}$,

(iii) $G(\sigma(t),\sigma(t),\lambda,\alpha) = [I + \mu(t)\mathcal{S}(t,\lambda)]G(t,\sigma(t),\lambda,\alpha) + \mathcal{J}$ *for every right-scattered point* $t \in [a,\infty)_{\mathbb{T}}$,

(iv) *for every* $t,s \in [a,\infty)_{\mathbb{T}}$ *such that* $t \notin \mathcal{T}(s)$, *the function* $G(\cdot,s,\lambda,\alpha)$ *solves the homogeneous system* $(\mathcal{S}_\lambda)$ *on the set* $\mathcal{T}(s)$, *where*

$$\mathcal{T}(s) := \left\{\tau \in [a,\infty)_{\mathbb{T}},\ \tau \neq \rho(s) \text{ if } s \text{ is left-scattered}\right\}, \tag{8.8}$$

(v) *the columns of* $G(\cdot,s,\lambda,\alpha)$ *belong to* $L^2_{\mathcal{W}}$ *for every* $s \in [a,\infty)_{\mathbb{T}}$, *and the columns of* $G(t,\cdot,\lambda,\alpha)$ *belong to* $L^2_{\mathcal{W}}$ *for every* $t \in [a,\infty)_{\mathbb{T}}$.

Proof. Condition (i) follows from the definition of $G(\cdot,s,\lambda,\alpha)$ in (8.5). Condition (ii) is a consequence of Lemma 8.2. Condition (iii) is proven from the definition of $G(\sigma(t),\sigma(t),\lambda,\alpha)$ in (8.5) by using Lemma 8.2 and $\widetilde{Z}(t,\lambda,\alpha) = \widetilde{Z}^\sigma(t,\lambda,\alpha) - \mu(t)\mathcal{S}(t,\lambda)\widetilde{Z}(t,\lambda,\alpha)$. Concerning condition (iv), the function $G(\cdot,s,\lambda,\alpha)$ solves the system $(\mathcal{S}_\lambda)$ on $[s,\infty)_{\mathbb{T}}$ because $\mathcal{X}_+(\cdot,\lambda,\alpha)$ solves this system on $[s,\infty)_{\mathbb{T}}$. If $s \in (a,\infty)_{\mathbb{T}}$ is left-dense, then $G(\cdot,s,\lambda,\alpha)$ solves $(\mathcal{S}_\lambda)$ on $[a,s)_{\mathbb{T}}$, since $\widetilde{Z}(\cdot,\lambda,\alpha)$ solves this system on $[a,s)_{\mathbb{T}}$. For the same reason $G(\cdot,s,\lambda,\alpha)$ solves $(\mathcal{S}_\lambda)$ on $[a,\rho(s))_{\mathbb{T}}$ if $s \in (a,\infty)_{\mathbb{T}}$ is left-scattered. Condition (v) follows from the definition of $G(\cdot,s,\lambda,\alpha)$ in (8.5) used with $t \geq s$ and from the fact that the columns of $\mathcal{X}_+(\cdot,\lambda,\alpha)$ belong to $L^2_{\mathcal{W}}$, by Theorem 6.7. The columns of $G(t,\cdot,\lambda,\alpha)$ then belong to $L^2_{\mathcal{W}}$ by part (i) of this lemma. □

Since by Lemma 8.3(v) the columns of $G(t,\cdot,\lambda,\alpha)$ belong to $L^2_{\mathcal{W}}$, the function

$$\widehat{z}(t,\lambda,\alpha) := -\int_a^\infty G(t,\sigma(s),\lambda,\alpha)\widetilde{\mathcal{W}}(s)f^\sigma(s)\Delta s, \quad t \in [a,\infty)_{\mathbb{T}}, \tag{8.9}$$

is well defined whenever $f \in L^2_{\mathcal{W}}$. Moreover, by using (8.6), we can write $\widehat{z}(t,\lambda,\alpha)$ as

$$\begin{aligned}\widehat{z}(t,\lambda,\alpha) = &-\mathcal{X}_+(t,\lambda,\alpha)\int_a^t \widetilde{Z}^{\sigma*}\left(s,\overline{\lambda},\alpha\right)\widetilde{\mathcal{W}}(s)f^\sigma(s)\Delta s\\ &-\widetilde{Z}(t,\lambda,\alpha)\int_t^\infty \mathcal{X}_+^{\sigma*}\left(s,\overline{\lambda},\alpha\right)\widetilde{\mathcal{W}}(s)f^\sigma(s)\Delta s, \quad t \in [a,\infty)_{\mathbb{T}}.\end{aligned} \tag{8.10}$$

Proposition 8.4. *For* $\alpha \in \Gamma$, $\lambda \in \mathbb{C} \setminus \mathbb{R}$, *and* $f \in L^2_{\mathcal{W}}$, *the function* $\widehat{z}(\cdot,\lambda,\alpha)$ *defined in* (8.9) *solves the nonhomogeneous system* (8.1) *with the initial condition* $\alpha\widehat{z}(a,\lambda,\alpha) = 0$.

Proof. By the time scales product rule (2.1) when we Δ-differentiate expression (8.10), we have for every $t \in [a,\infty)_{\mathbb{T}}$ (suppressing the dependence on α in the the following calculation)

$$\begin{aligned}
\hat{z}^{\Delta}(t,\lambda) &= -\mathcal{X}_{+}^{\Delta}(t,\lambda)\int_a^t \tilde{Z}^{\sigma*}\left(s,\bar{\lambda}\right)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s - \mathcal{X}_{+}^{\sigma}(t,\lambda)\tilde{Z}^{\sigma*}\left(t,\bar{\lambda}\right)\widetilde{\mathcal{W}}(t)f^{\sigma}(t) \\
&\quad - \tilde{Z}^{\Delta}(t,\lambda)\int_t^{\infty}\mathcal{X}_{+}^{\sigma*}\left(s,\bar{\lambda}\right)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s + \tilde{Z}^{\sigma}(t,\lambda)\mathcal{X}_{+}^{\sigma*}\left(t,\bar{\lambda}\right)\widetilde{\mathcal{W}}(t)f^{\sigma}(t) \\
&= \mathcal{S}(t,\lambda)\hat{z}(t,\lambda) - \left[\mathcal{X}_{+}^{\sigma}(t,\lambda)\tilde{Z}^{\sigma*}\left(t,\bar{\lambda}\right) - \tilde{Z}^{\sigma}(t,\lambda)\mathcal{X}_{+}^{\sigma*}\left(t,\bar{\lambda}\right)\right]\widetilde{\mathcal{W}}(t)f^{\sigma}(t) \\
&\overset{(8.7)}{=} \mathcal{S}(t,\lambda)\hat{z}(t,\lambda) - \mathcal{J}\widetilde{\mathcal{W}}(t)f^{\sigma}(t).
\end{aligned} \tag{8.11}$$

This shows that $\hat{z}(\cdot,\lambda,\alpha)$ is a solution of system (8.1). From (8.10) with $t = a$, we get

$$\alpha\hat{z}(a,\lambda,\alpha) = -\alpha\tilde{Z}(a,\lambda,\alpha)\int_a^{\infty}\mathcal{X}_{+}^{\sigma*}\left(s,\bar{\lambda},\alpha\right)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s = 0, \tag{8.12}$$

where we used the initial condition $\tilde{Z}(a,\lambda,\alpha) = -\mathcal{J}\alpha^*$ and $\alpha\mathcal{J}\alpha^* = 0$ coming from $\alpha \in \Gamma$. □

The following theorem provides further properties of the solution $\hat{z}(\cdot,\lambda,\alpha)$ of system (8.1). It is a generalization of [10, Lemma 4.2], [11, Theorem 7.5], [2, Theorem 5.2] to time scales.

Theorem 8.5. *Let $\alpha \in \Gamma$, $\lambda \in \mathbb{C}\setminus\mathbb{R}$, and $f \in L^2_{\mathcal{W}}$. Suppose that system $(\mathcal{S}_\lambda)$ is in the limit point or limit circle case. Then the solution $\hat{z}(\cdot,\lambda,\alpha)$ of system* (8.1) *defined in* (8.9) *belongs to $L^2_{\mathcal{W}}$ and satisfies*

$$\|\hat{z}(\cdot,\lambda,\alpha)\|_{\mathcal{W}} \le \frac{1}{|\mathrm{Im}(\lambda)|}\|f\|_{\mathcal{W}}, \tag{8.13}$$

$$\lim_{t\to\infty}\mathcal{X}_{+}^{*}(t,\nu,\alpha)\mathcal{J}\hat{z}(t,\lambda,\alpha) = 0, \quad \textit{for every } \nu \in \mathbb{C}\setminus\mathbb{R}. \tag{8.14}$$

Proof. To shorten the notation we suppress the dependence on α in all quantities appearing in this proof. Assume first that system $(\mathcal{S}_\lambda)$ is in the limit point case. For every $r \in [a,\infty)_{\mathbb{T}}$ we define the function $f_r(\cdot) := f(\cdot)$ on $[a,r]_{\mathbb{T}}$ and $f_r(\cdot) := 0$ on $(r,\infty)_{\mathbb{T}}$ and the function

$$\hat{z}_r(t,\lambda) := -\int_a^{\infty} G(t,\sigma(s),\lambda)\widetilde{\mathcal{W}}(s)f_r^{\sigma}(s)\Delta s = -\int_a^{r} G(t,\sigma(s),\lambda)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s. \tag{8.15}$$

For every $t \in [r,\infty)_{\mathbb{T}}$ we have as in (8.10) that

$$\hat{z}_r(t,\lambda) = -\mathcal{X}_{+}(t,\lambda)g(r,\lambda), \quad g(r,\lambda) := \int_a^r \tilde{Z}^{\sigma*}\left(s,\bar{\lambda}\right)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s. \tag{8.16}$$

to Contents, page 1

Since by Theorem 6.7 the solution $\mathcal{X}_+(\cdot,\lambda) \in L^2_{\mathcal{W}}$, (8.16) shows that $\hat{z}_r(\cdot,\lambda)$, being a multiple of $\mathcal{X}_+(\cdot,\lambda)$, also belongs to $L^2_{\mathcal{W}}$. Moreover, by Theorem 7.3,

$$\lim_{t\to\infty} \hat{z}_r^*(t,\lambda)\mathcal{J}\hat{z}_r(t,\lambda) \overset{(8.16)}{=} g^*(r,\lambda)\lim_{t\to\infty}\mathcal{X}_+^*(t,\lambda)\mathcal{J}\mathcal{X}_+(t,\lambda)g(r,\lambda) \overset{(7.4)}{=} 0. \tag{8.17}$$

On the other hand, $\hat{z}_r^*(a,\lambda)\mathcal{J}\hat{z}_r(a,\lambda) = 0$, and for any $t \in [a,\infty)_{\mathbb{T}}$ identity (8.3) implies

$$\begin{aligned} &\hat{z}_r^*(t,\lambda)\mathcal{J}\hat{z}_r(t,\lambda) \\ &\quad = -2i\,\mathrm{Im}(\lambda)\int_a^t \hat{z}_r^{\sigma*}(s,\lambda)\widetilde{\mathcal{W}}(s)\hat{z}_r^{\sigma}(s,\lambda)\Delta s + 2i\,\mathrm{Im}\left(\int_a^t \hat{z}_r^{\sigma*}(s,\lambda)\widetilde{\mathcal{W}}(s)f_r^{\sigma}(s)\Delta s\right). \end{aligned} \tag{8.18}$$

Combining (8.18), where $t \to \infty$, formula (8.17), and the definition on $f_r(\cdot)$ yields

$$\|\hat{z}_r(\cdot,\lambda)\|^2_{\mathcal{W}} = \int_a^\infty \hat{z}_r^{\sigma*}(s,\lambda)\widetilde{\mathcal{W}}(s)\hat{z}_r^{\sigma}(s,\lambda)\Delta s = \frac{1}{\mathrm{Im}(\lambda)}\mathrm{Im}\left(\int_a^r \hat{z}_r^{\sigma*}(s,\lambda)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s\right). \tag{8.19}$$

By using the Cauchy-Schwarz inequality (C-S) on time scales and $\widetilde{\mathcal{W}}(\cdot) \geq 0$, we then have

$$\begin{aligned} \|\hat{z}_r(\cdot,\lambda)\|^2_{\mathcal{W}} &= \frac{1}{2i\,\mathrm{Im}(\lambda)}\left[\int_a^r \hat{z}_r^{\sigma*}(s,\lambda)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s - \int_a^r f^{\sigma*}(s)\widetilde{\mathcal{W}}(s)\hat{z}_r^{\sigma}(s,\lambda)\Delta s\right] \\ &\leq \frac{1}{|\mathrm{Im}(\lambda)|}\left|\int_a^r \hat{z}_r^{\sigma*}(s,\lambda)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s\right| \\ &\overset{\text{C-S}}{\leq} \frac{1}{|\mathrm{Im}(\lambda)|}\left(\int_a^r \hat{z}_r^{\sigma*}(s,\lambda)\widetilde{\mathcal{W}}(s)\hat{z}_r^{\sigma}(s,\lambda)\Delta s\right)^{1/2}\left(\int_a^r f^{\sigma*}(s)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s\right)^{1/2} \\ &\leq \frac{1}{|\mathrm{Im}(\lambda)|}\|\hat{z}_r(\cdot,\lambda)\|_{\mathcal{W}}\|f\|_{\mathcal{W}}. \end{aligned} \tag{8.20}$$

Since $\|\hat{z}_r(\cdot,\lambda)\|_{\mathcal{W}}$ is finite by $\hat{z}_r(\cdot,\lambda) \in L^2_{\mathcal{W}}$, we get from the above calculation that

$$\|\hat{z}_r(\cdot,\lambda)\|_{\mathcal{W}} \leq \frac{1}{|\mathrm{Im}(\lambda)|}\|f\|_{\mathcal{W}}. \tag{8.21}$$

We will prove that (8.21) implies estimate (8.13) by the convergence argument. For any $t, r \in [a,\infty)_{\mathbb{T}}$ we observe that

$$\hat{z}(t,\lambda) - \hat{z}_r(t,\lambda) = -\int_r^\infty G(t,\sigma(s),\lambda)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s. \tag{8.22}$$

Now we fix $q \in [a,r)_{\mathbb{T}}$. By the definition of $G(\cdot,\cdot,\lambda)$ in (8.5) we have for every $t \in [a,q]_{\mathbb{T}}$

$$\hat{z}(t,\lambda) - \hat{z}_r(t,\lambda) = -\widetilde{Z}(t,\lambda)\int_r^\infty \mathcal{X}_+^*\left(\sigma(s),\overline{\lambda}\right)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s. \tag{8.23}$$

to Contents, page 1

Since the functions $\mathcal{X}_+(\cdot,\overline{\lambda})$ and $f(\cdot)$ belong to $L^2_{\mathcal{W}}$, it follows that the right-hand side of (8.23) converges to zero as $r \to \infty$ for every $t \in [a,q]_{\mathbb{T}}$. Hence, $\widehat{z}_r(\cdot,\lambda)$ converges to the function $\widehat{z}(\cdot,\lambda)$ uniformly on $[a,q]_{\mathbb{T}}$. Since $\widehat{z}(\cdot,\lambda) = \widehat{z}_r(\cdot,\lambda)$ on $[a,q]_{\mathbb{T}}$, we have by $\widetilde{\mathcal{W}}(\cdot) \geq 0$ and (8.21) that

$$\int_a^q \widehat{z}^{\sigma *}(s,\lambda)\widetilde{\mathcal{W}}(s)\widehat{z}^{\sigma}(s,\lambda)\Delta s \leq \|\widehat{z}_r(\cdot,\lambda)\|^2_{\mathcal{W}} \overset{(8.21)}{\leq} \frac{1}{|\operatorname{Im}(\lambda)|^2}\|f\|^2_{\mathcal{W}}. \tag{8.24}$$

Since $q \in [a,\infty)_{\mathbb{T}}$ was arbitrary, inequality (8.24) implies the result in (8.13). In the limit circle case inequality (8.13) follows by the same argument by using the fact that all solutions of system $(\mathcal{S}_\lambda)$ belong to $L^2_{\mathcal{W}}$.

Now we prove the existence of the limit (8.14). Assume that the system $(\mathcal{S}_\lambda)$ is in the limit point case, and let $\nu \in \mathbb{C} \setminus \mathbb{R}$ be arbitrary. Following the argument in the proof of [30, Lemma 4.1] and [2, Theorem 5.2], we have from identity (8.3) that for any $r,t \in [a,\infty)_{\mathbb{T}}$

$$\begin{aligned}\mathcal{X}_+^*(t,\nu)\mathcal{J}\widehat{z}_r(t,\lambda) = \mathcal{X}_+^*(a,\nu)\mathcal{J}\widehat{z}_r(a,\lambda) + (\overline{\nu}-\lambda)\int_a^t \mathcal{X}_+^{\sigma *}(s,\nu)\widetilde{\mathcal{W}}(s)\widehat{z}_r^{\sigma}(s,\lambda)\Delta s \\ + \int_a^t \mathcal{X}_+^{\sigma *}(s,\nu)\widetilde{\mathcal{W}}(s)f_r^{\sigma}(s)\Delta s.\end{aligned} \tag{8.25}$$

Since for $t \in [r,\infty)_{\mathbb{T}}$ equality (8.16) holds, it follows that

$$\lim_{t\to\infty}\mathcal{X}_+^*(t,\nu)\mathcal{J}\widehat{z}_r(t,\lambda) = -\lim_{t\to\infty}\mathcal{X}_+^*(t,\nu)\mathcal{J}\mathcal{X}_+(t,\lambda)g(r,\lambda) \overset{(7.4)}{=} 0. \tag{8.26}$$

Hence, by (8.25),

$$\mathcal{X}_+^*(a,\nu)\mathcal{J}\widehat{z}_r(a,\lambda) = (\lambda-\overline{\nu})\int_a^\infty \mathcal{X}_+^{\sigma *}(s,\nu)\widetilde{\mathcal{W}}(s)\widehat{z}_r^{\sigma}(s,\lambda)\Delta s - \int_a^r \mathcal{X}_+^{\sigma *}(s,\nu)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s. \tag{8.27}$$

By the uniform convergence of $\widehat{z}_r(\cdot,\lambda)$ to $\widehat{z}(\cdot,\lambda)$ on compact intervals, we get from (8.27) with $r \to \infty$ the equality

$$\mathcal{X}_+^*(a,\nu)\mathcal{J}\widehat{z}(a,\lambda) = (\lambda-\overline{\nu})\int_a^\infty \mathcal{X}_+^{\sigma *}(s,\nu)\widetilde{\mathcal{W}}(s)\widehat{z}^{\sigma}(s,\lambda)\Delta s - \int_a^\infty \mathcal{X}_+^{\sigma *}(s,\nu)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s. \tag{8.28}$$

On the other hand, by (8.3), we obtain for every $t \in [a,\infty)_{\mathbb{T}}$

$$\begin{aligned}\mathcal{X}_+^*(t,\nu)\mathcal{J}\widehat{z}(t,\lambda) = \mathcal{X}_+^*(a,\nu)\mathcal{J}\widehat{z}(a,\lambda) + (\overline{\nu}-\lambda)\int_a^t \mathcal{X}_+^{\sigma *}(s,\nu)\widetilde{\mathcal{W}}(s)\widehat{z}^{\sigma}(s,\lambda)\Delta s \\ + \int_a^t \mathcal{X}_+^{\sigma *}(s,\nu)\widetilde{\mathcal{W}}(s)f^{\sigma}(s)\Delta s.\end{aligned} \tag{8.29}$$

to Contents, page 1

Upon taking the limit in (8.29) as $t \to \infty$ and using equality (8.28), we conclude that the limit in (8.14) holds true.

In the limit circle case, the limit in (8.14) can be proved similarly as above, because all the solutions of system $(\mathcal{S}_\lambda)$ now belong to $L^2_{\mathcal{W}}$. However, in this case, we can apply a direct argument to show that (8.14) holds. By formula (8.10) we get for every $t \in [a,\infty)_{\mathbb{T}}$

$$\begin{aligned}\mathcal{X}_+^*(t,\nu)\mathcal{J}\widehat{z}(t,\lambda) = -\mathcal{X}_+^*(t,\nu)\mathcal{J}\mathcal{X}_+(t,\lambda)\int_a^t \widetilde{Z}^{\sigma *}\left(s,\overline{\lambda}\right)\widetilde{\mathcal{W}}(s)f^\sigma(s)\Delta s \\ - \mathcal{X}_+^*(t,\nu)\mathcal{J}\widetilde{Z}(t,\lambda)\int_t^\infty \mathcal{X}_+^{\sigma *}\left(s,\overline{\lambda}\right)\widetilde{\mathcal{W}}(s)f^\sigma(s)\Delta s.\end{aligned} \tag{8.30}$$

The limit of the first term in (8.30) is zero because $\mathcal{X}_+^*(t,\nu)\mathcal{J}\mathcal{X}_+(t,\lambda)$ tends to zero for $t \to \infty$ by (7.4), and it is multiplied by a convergent integral as $t \to \infty$. Since the columns of $\widetilde{Z}(\cdot,\lambda)$ belong to $L^2_{\mathcal{W}}$, the function $\mathcal{X}_+^*(\cdot,\nu)\mathcal{J}\widetilde{Z}(\cdot,\lambda)$ is bounded on $[a,\infty)_{\mathbb{T}}$, and it is multiplied by an integral converging to zero as $t \to \infty$. Therefore, formula (8.14) follows. □

In the last result of this paper we construct another solution of the nonhomogeneous system (8.1) satisfying condition (8.14) and such that it starts with a possibly nonzero initial condition at $t = a$. It can be considered as an extension of Theorem 8.5.

Corollary 8.6. *Let $\alpha \in \Gamma$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Assume that $(\mathcal{S}_\lambda)$ is in the limit point or limit circle case. For $f \in L^2_{\mathcal{W}}$ and $v \in \mathbb{C}^n$ we define*

$$\widetilde{z}(t,\lambda,\alpha) := \mathcal{X}_+(t,\lambda,\alpha)v + \widehat{z}(t,\lambda,\alpha), \quad \forall t \in [a,\infty)_{\mathbb{T}}, \tag{8.31}$$

where $\widehat{z}(\cdot,\lambda,\alpha)$ is given in (8.9). *Then $\widetilde{z}(\cdot,\lambda,\alpha)$ solves the system* (8.1) *with $\alpha\widetilde{z}(a,\lambda,\alpha) = v$,*

$$\|\widetilde{z}(\cdot,\lambda,\alpha)\|_{\mathcal{W}} \le \frac{1}{|\mathrm{Im}(\lambda)|}\|f\|_{\mathcal{W}} + \|\mathcal{X}_+(\cdot,\lambda,\alpha)v\|_{\mathcal{W}}, \tag{8.32}$$

$$\lim_{t\to\infty}\mathcal{X}_+^*(t,\nu,\alpha)\mathcal{J}\widetilde{z}(t,\lambda,\alpha) = 0, \quad \textit{for every } \nu \in \mathbb{C}\setminus\mathbb{R}. \tag{8.33}$$

In addition, if the system $(\mathcal{S}_\lambda)$ is in the limit point case, then $\widetilde{z}(\cdot,\lambda,\alpha)$ is the only $L^2_{\mathcal{W}}$ solution of (8.1) *satisfying $\alpha\widetilde{z}(a,\lambda,\alpha) = v$.*

Proof. As in the previous proof we suppress the dependence on α. Since the function $\mathcal{X}_+(\cdot,\lambda)v$ solves $(\mathcal{S}_\lambda)$, it follows from Proposition 8.4 that $\widetilde{z}(\cdot,\lambda,\alpha)$ solves the system (8.1) and $\alpha\widetilde{z}(a,\lambda) = \alpha\mathcal{X}_+(a,\lambda)v = v$. Next, $\widetilde{z}(\cdot,\lambda) \in L^2_{\mathcal{W}}$ as a sum of two $L^2_{\mathcal{W}}$ functions. The limit in (8.33) follows from the limit (8.14) of Theorem 8.5 and from identity (7.4), because

$$\lim_{t\to\infty}\mathcal{X}_+^*(t,\nu)\mathcal{J}\widetilde{z}(t,\lambda) = \lim_{t\to\infty}\{\mathcal{X}_+^*(t,\nu)\mathcal{J}\mathcal{X}_+(t,\lambda)v + \mathcal{X}_+^*(t,\nu)\mathcal{J}\widehat{z}(t,\lambda)\} = 0. \tag{8.34}$$

Inequality (8.32) is obtained from estimate (8.13) by the triangle inequality.

to Contents, page 1

Now we prove the uniqueness of $\tilde{z}(\cdot,\lambda)$ in the case of $(\mathcal{S}_\lambda)$ being of the limit point type. If $z_1(\cdot,\lambda)$ and $z_2(\cdot,\lambda)$ are two $L^2_{\mathcal{W}}$ solutions of (8.1) satisfying $\alpha z_1(a,\lambda) = v = \alpha z_2(a,\lambda)$, then their difference $z(\cdot,\lambda) := z_1(\cdot,\lambda) - z_2(\cdot,\lambda)$ also belongs to $L^2_{\mathcal{W}}$ and solves system $(\mathcal{S}_\lambda)$ with $\alpha z(\cdot,\lambda) = 0$. Since $z(\cdot,\lambda) = \Psi(\cdot,\lambda)c$ for some $c \in \mathbb{C}^{2n}$, the initial condition $\alpha z(\cdot,\lambda) = 0$ implies through (4.7) that $z(\cdot,\lambda) = \tilde{Z}(\cdot,\lambda)d$ for some $d \in \mathbb{C}^n$. If $d \neq 0$, then $z(\cdot,\lambda) \notin L^2_{\mathcal{W}}$, because in the limit point case the columns of $\tilde{Z}(\cdot,\lambda)$ do not belong to $L^2_{\mathcal{W}}$, which is a contradiction. Therefore, $d = 0$ and the uniqueness of $\tilde{z}(\cdot,\lambda)$ is established. □

Acknowledgments

The research was supported by the Czech Science Foundation under Grant 201/09/J009, by the research project MSM 0021622409 of the Ministry of Education, Youth, and Sports of the Czech Republic, and by the Grant MUNI/A/0964/2009 of Masaryk University.

References

[1] M. Bohner and S. Sun, "Weyl-Titchmarsh theory for symplectic difference systems," *Applied Mathematics and Computation*, vol. 216, no. 10, pp. 2855–2864, 2010.

[2] S. Clark and P. Zemánek, "On a Weyl-Titchmarsh theory for discrete symplectic systems on a half line," *Applied Mathematics and Computation*, vol. 217, no. 7, pp. 2952–2976, 2010.

[3] S. Sun, M. Bohner, and S. Chen, "Weyl-Titchmarsh theory for Hamiltonian dynamic systems," *Abstract and Applied Analysis*, vol. 2010, Article ID 514760, 18 pages, 2010.

[4] C. Zhang and L. Zhang, "Classification for a class of second-order singular equations on time scales," in *Proceedings of the 8th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (Qingdao, China, 2007)*, W. Feng and F. Gao, Eds., vol. 3, pp. 73–76, IEEE, Los Alamitos, Calif, USA, 2007.

[5] D. R. Anderson, "Titchmarsh-Sims-Weyl theory for complex Hamiltonian systems on Sturmian time scales," *Journal of Mathematical Analysis and Applications*, vol. 373, no. 2, pp. 709–725, 2011.

[6] B. M. Brown, W. D. Evans, and M. Plum, "Titchmarsh-Sims-Weyl theory for complex Hamiltonian systems," *Proceedings of the London Mathematical Society*, vol. 87, no. 2, pp. 419–450, 2003.

[7] D. Chelkak and E. Korotyaev, "Weyl-Titchmarsh functions of vector-valued Sturm-Liouville operators on the unit interval," *Journal of Functional Analysis*, vol. 257, no. 5, pp. 1546–1588, 2009.

[8] S. Clark and F. Gesztesy, "Weyl-Titchmarsh M-function asymptotics, local uniqueness results, trace formulas, and Borg-type theorems for Dirac operators," *Transactions of the American Mathematical Society*, vol. 354, no. 9, pp. 3475–3534, 2002.

[9] S. Clark and F. Gesztesy, "On Weyl-Titchmarsh theory for singular finite difference Hamiltonian systems," *Journal of Computational and Applied Mathematics*, vol. 171, no. 1-2, pp. 151–184, 2004.

[10] D. B. Hinton and J. K. Shaw, "Hamiltonian systems of limit point or limit circle type with both endpoints singular," *Journal of Differential Equations*, vol. 50, no. 3, pp. 444–464, 1983.

[11] A. M. Krall, "$M(\lambda)$ theory for singular Hamiltonian systems with one singular point," *SIAM Journal on Mathematical Analysis*, vol. 20, no. 3, pp. 664–700, 1989.

[12] S. J. Monaquel and K. M. Schmidt, "On M-functions and operator theory for non-self-adjoint discrete Hamiltonian systems," *Journal of Computational and Applied Mathematics*, vol. 208, no. 1, pp. 82–101, 2007.

[13] G. Ren and Y. Shi, "The defect index of singular symmetric linear difference equations with real coefficients," *Proceedings of the American Mathematical Society*, vol. 138, no. 7, pp. 2463–2475, 2010.

[14] Y. Shi, "Weyl-Titchmarsh theory for a class of discrete linear Hamiltonian systems," *Linear Algebra and Its Applications*, vol. 416, no. 2-3, pp. 452–519, 2006.

[15] H. Weyl, "Über gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen," *Mathematische Annalen*, vol. 68, no. 2, pp. 220–269, 1910.

[16] E. C. Titchmarsh, "On expansions in eigenfunctions," *Journal of the London Mathematical Society*, vol. 14, pp. 274–278, 1939.

to Contents, page 1

[17] E. C. Titchmarsh, *Eigenfunction Expansions Associated with Second-Order Differential Equations. Part I*, Clarendon Press, Oxford, UK, 2nd edition, 1962.
[18] C. Bennewitz and W. N. Everitt, "The Titchmarsh-Weyl eigenfunction expansion theorem for Sturm-Liouville differential equations," in *Sturm-Liouville Theory. Past and Presen, Proceedings of the International Colloquium (Geneva, 2003)*, W. O. Amrein, A. M. Hinz, and D. B. Pearson, Eds., pp. 137–171, Birkhäuser, Basel, Switzerland, 2005.
[19] W. N. Everitt, "A personal history of the m-coefficient," *Journal of Computational and Applied Mathematics*, vol. 171, no. 1-2, pp. 185–197, 2004.
[20] J. Weidmann, "Spectral theory of Sturm-Liouville operators approximation by regular problems," in *Sturm-Liouville Theory. Past and Present, , Proceedings of the International Colloquium (Geneva, 2003)*, W. O. Amrein, A. M. Hinz, and D. B. Pearson, Eds., pp. 75–98, Birkhäuser, Basel, Switzerland, 2005.
[21] A. R. Sims, "Secondary conditions for linear differential operators of the second order," *Journal of Mathematics and Mechanics*, vol. 6, pp. 247–285, 1957.
[22] F. V. Atkinson, *Discrete and Continuous Boundary Problems*, vol. 8 of *Mathematics in Science and Engineering*, Academic Press, New York, NY, USA, 1964.
[23] S. Clark and F. Gesztesy, "Weyl-Titchmarsh M-function asymptotics for matrix-valued Schrödinger operators," *Proceedings of the London Mathematical Society*, vol. 82, no. 3, pp. 701–724, 2001.
[24] W. N. Everitt, D. B. Hinton, and J. K. Shaw, "The asymptotic form of the Titchmarsh-Weyl coefficient for Dirac systems," *Journal of the London Mathematical Society*, vol. 27, no. 3, pp. 465–476, 1983.
[25] D. B. Hinton and A. Schneider, "On the Titchmarsh-Weyl coefficients for singular S-Hermitian systems. I," *Mathematische Nachrichten*, vol. 163, pp. 323–342, 1993.
[26] D. B. Hinton and A. Schneider, "On the Titchmarsh-Weyl coefficients for singular S-Hermitian systems. II," *Mathematische Nachrichten*, vol. 185, pp. 67–84, 1997.
[27] D. B. Hinton and J. K. Shaw, "On Titchmarsh-Weyl $M(\lambda)$-functions for linear Hamiltonian systems," *Journal of Differential Equations*, vol. 40, no. 3, pp. 316–342, 1981.
[28] D. B. Hinton and J. K. Shaw, "On the spectrum of a singular Hamiltonian system," *Quaestiones Mathematicae*, vol. 5, no. 1, pp. 29–81, 1983.
[29] D. B. Hinton and J. K. Shaw, "Parameterization of the $M(\lambda)$ function for a Hamiltonian system of limit circle type," *Proceedings of the Royal Society of Edinburgh. Section A*, vol. 93, no. 3-4, pp. 349–360, 1983.
[30] D. B. Hinton and J. K. Shaw, "On boundary value problems for Hamiltonian systems with two singular points," *SIAM Journal on Mathematical Analysis*, vol. 15, no. 2, pp. 272–286, 1984.
[31] D. B. Hinton and J. K. Shaw, "On the spectrum of a singular Hamiltonian system. II," *Quaestiones Mathematicae*, vol. 10, no. 1, pp. 1–48, 1986.
[32] V. I. Kogan and F. S. Rofe-Beketov, "On square-integrable solutions of symmetric systems of differential equations of arbitrary order," *Proceedings of the Royal Society of Edinburgh. Section A*, vol. 74, pp. 5–40, 1976.
[33] A. M. Krall, "$M(\lambda)$ theory for singular Hamiltonian systems with two singular points," *SIAM Journal on Mathematical Analysis*, vol. 20, no. 3, pp. 701–715, 1989.
[34] A. M. Krall, "A limit-point criterion for linear Hamiltonian systems," *Applicable Analysis*, vol. 61, no. 1-2, pp. 115–119, 1996.
[35] M. Lesch and M. Malamud, "On the deficiency indices and self-adjointness of symmetric Hamiltonian systems," *Journal of Differential Equations*, vol. 189, no. 2, pp. 556–615, 2003.
[36] J. Qi and S. Chen, "Strong limit-point classification of singular Hamiltonian expressions," *Proceedings of the American Mathematical Society*, vol. 132, no. 6, pp. 1667–1674, 2004.
[37] C. Remling, "Geometric characterization of singular self-adjoint boundary conditions for Hamiltonian systems," *Applicable Analysis*, vol. 60, no. 1-2, pp. 49–61, 1996.
[38] Y. Shi, "On the rank of the matrix radius of the limiting set for a singular linear Hamiltonian system," *Linear Algebra and Its Applications*, vol. 376, pp. 109–123, 2004.
[39] R. Nevanlinna, "Asymptotische Entwicklungen beschränkter Funktionen und das Stieltjessche Momenten problem," *Annales Academiæ Scientiarum Fennicæ A*, vol. 18, no. 5, p. 53, 1922 (German).
[40] D. B. Hinton and R. T. Lewis, "Spectral analysis of second order difference equations," *Journal of Mathematical Analysis and Applications*, vol. 63, no. 2, pp. 421–438, 1978.
[41] A. B. Mingarelli, "A limit-point criterion for a three-term recurrence relation," *Comptes Rendus Mathématiques de l'Académie des Sciences*, vol. 3, no. 3, pp. 171–175, 1981.
[42] H. Sun and Y. Shi, "Strong limit point criteria for a class of singular discrete linear Hamiltonian systems," *Journal of Mathematical Analysis and Applications*, vol. 336, no. 1, pp. 224–242, 2007.
[43] J. Weiss, "Limit-point criteria for a second order dynamic equation on time scales," *Nonlinear Dynamics and Systems Theory of Research and Surveys*, vol. 9, no. 1, pp. 99–108, 2009.

to Contents, page 1

[44] W. Kratz, R. Šimon Hilscher, and V. Zeidan, "Eigenvalue and oscillation theorems for time scale symplectic systems," *International Journal of Dynamical Systems and Differential Equations*, vol. 3, no. 1-2, pp. 84–131, 2011.

[45] C. D. Ahlbrandt, M. Bohner, and J. Ridenhour, "Hamiltonian systems on time scales," *Journal of Mathematical Analysis and Applications*, vol. 250, no. 2, pp. 561–578, 2000.

[46] O. Došlý, S. Hilger, and R. Hilscher, "Symplectic dynamic systems," in *Advances in Dynamic Equations on Time Scales*, M. Bohner and A. Peterson, Eds., pp. 293–334, Birkhäuser, Boston, Mass, USA, 2003.

[47] R. Hilscher and V. Zeidan, "Time scale symplectic systems without normality," *Journal of Differential Equations*, vol. 230, no. 1, pp. 140–173, 2006.

[48] R. Hilscher and V. Zeidan, "Applications of time scale symplectic systems without normality," *Journal of Mathematical Analysis and Applications*, vol. 340, no. 1, pp. 451–465, 2008.

[49] R. Hilscher and V. Zeidan, "Riccati equations for abnormal time scale quadratic functionals," *Journal of Differential Equations*, vol. 244, no. 6, pp. 1410–1447, 2008.

[50] R. Hilscher and V. Zeidan, "Reid roundabout theorems for time scale symplectic systems," in *"Discrete Dynamics and Difference Equations", Proceedings of the 12th International Conference on Difference Equations and Applications (Lisbon, 2007)*, S. Elaydi, H. Oliveira, J. M. Ferreira, and J. F. Alves, Eds., World Scientific Publishing, London, UK, 2010.

[51] R. Hilscher and P. Zemánek, "Trigonometric and hyperbolic systems on time scales," *Dynamic Systems and Applications*, vol. 18, no. 3-4, pp. 483–505, 2009.

[52] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales: An Introduction with Applications*, Birkhäuser, Boston, Mass, USA, 2001.

[53] S. Hilger, "Analysis on measure chains—a unified approach to continuous and discrete calculus," *Results in Mathematics*, vol. 18, no. 1-2, pp. 18–56, 1990.

[54] R. Hilscher and V. Zeidan, "Time scale embedding theorem and coercivity of quadratic functionals," *Analysis*, vol. 28, no. 1, pp. 1–28, 2008.

[55] F. R. Gantmacher, *The Theory of Matrices. Vol. 1*, AMS Chelsea Publishing, Providence, RI, USA, 1998, translated from the Russian by K. A. Hirsch. Reprint of the 1959 translation.

[56] D. S. Bernstein, *Matrix Mathematics. Theory, Facts, and Formulas with Application to Linear Systems Theory*, Princeton University Press, Princeton, NJ, USA, 2005.

[57] O. Došlý and R. Hilscher, "Disconjugacy, transformations and quadratic functionals for symplectic dynamic systems on time scales," *Journal of Difference Equations and Applications*, vol. 7, no. 2, pp. 265–295, 2001.

[58] C. D. Ahlbrandt and A. C. Peterson, *Discrete Hamiltonian Systems: Difference Equations, Continued Fractions, and Riccati Equations*, vol. 16 of *Kluwer Texts in the Mathematical Sciences*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

[59] M. Bohner and O. Došlý, "Disconjugacy and transformations for symplectic systems," *The Rocky Mountain Journal of Mathematics*, vol. 27, no. 3, pp. 707–743, 1997.

[60] W. Kratz, *Quadratic Functionals in Variational Analysis and Control Theory*, vol. 6 of *Mathematical Topics*, Akademie, Berlin, Germany, 1995.

[61] P. Piccione and D. V. Tausk, "Index theorems for symplectic systems," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 47, no. 5, pp. 3031–3046, 2001.

[62] R. P. Agarwal, M. Bohner, and P. J. Y. Wong, "Sturm-Liouville eigenvalue problems on time scales," *Applied Mathematics and Computation*, vol. 99, no. 2-3, pp. 153–166, 1999.

[63] M. Bohner and R. Hilscher, "An eigenvalue problem for linear Hamiltonian dynamic systems," *Fasciculi Mathematici*, no. 35, pp. 35–49, 2005.

[64] W. T. Reid, *Ordinary Differential Equations*, John Wiley & Sons, New York, NY, USA, 1971.

[65] R. Hilscher, V. Zeidan, and W. Kratz, "Differentiation of solutions of dynamic equations on time scales with respect to parameters," *Advances in Dynamical Systems and Applications*, vol. 4, no. 1, pp. 35–54, 2009.

[66] S. A. Orlov, "Nested matrix discs that depend analytically on a parameter, and theorems on the invariance of the ranks of the radii of the limit matrix discs," *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, vol. 40, no. 3, pp. 593–644, 710, 1976 (Russian), translated in *Mathematics of the USSR-Izvestiya*, vol. 10, pp. 565–613, 1976.

[67] Z. Zheng, J. Qi, and S. Chen, "Eigenvalues below the lower bound of minimal operators of singular Hamiltonian expressions," *Computers & Mathematics with Applications*, vol. 56, no. 11, pp. 2825–2833, 2008.

[68] F. M. Atici and G. S. Guseinov, "On Green's functions and positive solutions for boundary value problems on time scales," *Journal of Computational and Applied Mathematics*, vol. 141, no. 1-2, pp. 75–99, 2002.

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.9-16

PACS numbers: 42.62.Cf, 73.61.At

LASER SCRIBING OPTIMIZATION OF RF MAGNETRON SPUTTERED MOLYBDENUM THIN FILMS

***V.B. Padhiar*[1], *A.L. Patel*[2#], *J.R. Ray*[1], *M.S. Desai*[1], *C.J. Panchal*[1*]**

[1] Department of Applied Physics, M.S. University of Baroda,
Vadodara – 390001, Gujarat India
[*] E-mail: cjpanchal_msu@yahoo.com

[2] Director, Sahjanand Laser Technology Ltd.,
E-30, G. I.D.C. Electronics Estate, Sector 26,
Ghandhinagar-382 028, Gujarat, India
[#] E-mail: arvindp@sahajanandlaser.com

The optimization process of laser scribing of back contacts is carried out by varying different parameters of laser and thickness of Molybdenum (Mo) thin-films. Mo thin films were deposited by RF magnetron sputtering on the organically cleaned soda lime glass substrate. The thickness of Mo was in the range of 60 *nm to* 800 *nm. For the scribing process the laser power and the laser pulse frequency were varied. Different thickness of Mo shows the different scribe behavior. The optimized process provides a successful isolative laser scribing, having a minimum scribe line width, of Mo layer on glass substrate without any presence of walls, ridges, or collars in scribed areas.*

Keywords: *LASER SCRIBING, MOLYBDENUM THIN FILM, LASER POWER, PULSE FREQUENCY, SCRIBE LINE WIDTH.*

(Received 04 February 2011)

1. INTRODUCTION

Thin film solar cell technology provides a best solution towards more economic solar cell energy. The first step in manufacture of thin film solar cell is the deposition of molybdenum (Mo) back contact layer on soda lime glass substrate and the next step is the isolative laser scribing of this layer for making a metal back contacts between the layers. However, the problems are arising with a heat-affected zone around the scribed area due to the laser pulse beam [1]. The scribing should be done in such a way so that the scribed line should not contain bridges or process residue in the grooves that could cause an electrical connection and short-circuit the cell [2]. To increase the solar cell efficiency, the scribing of Mo layers desires minimum scribed area to have maximum active solar cell area [3] and so, minimize the dead area in the cell. Optimum operation for thin-film PV materials has been investigated by several PV manufacturers [4, 5, 6].

In this study, we present the results regarding the optimization of isolative laser scribing of sputtered-Mo thin-film deposited on soda lime glass substrate. The optimization process is performed by varying different laser parameters and thickness of the Mo thin-film in order to achieve lowest resistivity films. The result of successful scribing yields reliable, reproducible clean scribes without any presence of the buckling, ridges, or collars in the scribed areas.

to Contents, page 1

2. EXPERIMENTAL

The processes of laser scribing of Mo thin-films were done using commercially available the laser system that has the multi diode pumped fiber laser (20 W average maximum power) [Model No. Akhsar Fiber-Pro, Sahajanand Laser Technology Ltd., INDIA] Mo thin films used for the scribing have a different thickness from 60 nm to 800 nm. Mo thin films were deposited by RF magnetron sputtering at 1mTorr working pressure and 100 W RF power. The electrical property viz. sheet resistance of the Mo thin films was measured using four point probe method. The laser system, which was used for the scribing of Mo thin film, has a specification shown in Table 1.

Table 1 *– Laser system parameters used for the scribing the Mo thin films*

Laser	Multi Diode Pump Fiber Laser
Nominal average power	20 W (optional 10 W)
Maximum peak power	> 7.5 kW
Power tunability	10 - 100 %
Pulse repetation rate	20 - 80 kHz
Wavelength	1060 ± 10 nm
Pulse duration@20 kHz	<120 ns
Class	IV
Power stability	> 95 %
Pulse energy@20 kHz	1 mJ
Beam quality	1.5 (M^2)
Output beam diameter($1/e^2$)	9 mm
Scribe speed	upto 1000 mm/s
Inbuilt guide (marking) laser	He-Ne laser (660 nm & 0.5 mW)

In our study we have kept the scribing speed constant at 500 mm/sec in order to achieve lowest scribed width. Scribing process of Mo thin film was optimized by varying both the laser power and the pulse frequency simultaneously. First keeping the maximum average power 20 W and varying the pulse frequency from 1 to 80 Hz to get the optimum pulse frequency and by using that, vary the average power up to minimum of 1 W. Different thickness of the Mo shows different kind of scribe pattern. The smoothness of the scribed was observed using a polarization microscope (LABOURLUX 11, Leitz).

3. RESULTS AND DISCUSSIONS

Mo thin films having different thickness (60 nm to 800 nm) shows a different electrical conductivity [7]. The measured sheet resistance of the Mo thin films as a function of thickness of the films is shown in Fig. 1. Different thickness of the Mo thin films shows different kind of the scribing behavior. Ablation with a train of laser pulses per spot and scribing lines were performed on molybdenum back-contact from film side. The laser spot overlap along a scanning line was controlled by the translation speed at a constant pulse repetition rate.

to Contents, page 1

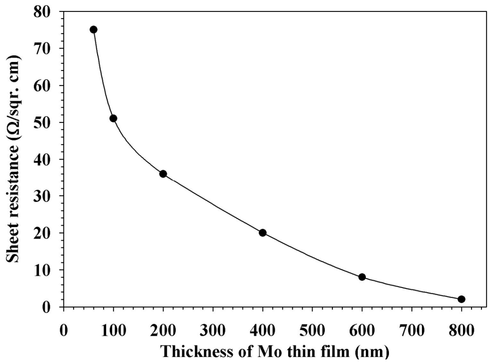


Fig. 1 *– Variation in the sheet resistance of Mo thin films as a function of thickness indicates that higher thickness shows higher conductivity*

Various combinations of laser power and pulse frequency were used for selective ablation of the films. Optimal regimes for laser processing of each layer were estimated depending on the Mo film thickness. The beam overlap or the number of pulses played an important role in the processing selectivity because the ablation threshold was sensitive to accumulation of the irradiation dose.

3.1 Effect of laser power and pulse frequency

However, at certain scribe parameters, namely the laser power together with laser scribe speed and pulse frequency, the scribes were superior to the high-energy scribes in two immediate respects: 1) The scribe profile had sharply defined edges without significant formation of walls, and 2) the glass remained undamaged, glass damage invariably appearing at higher energy densities.

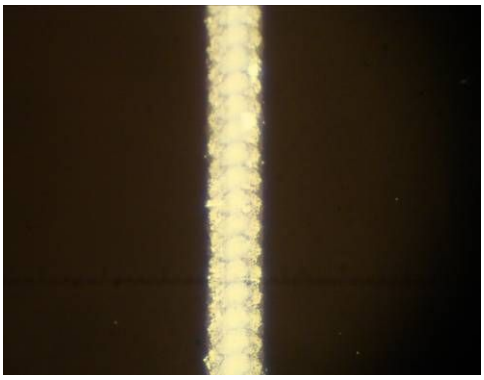


Fig. 2 *– Sharp scribed line (140 μm) of Mo thin-film has 2 watt laser power and 1 Hz pulse frequency*

The scribing image (Fig. 2) of Mo thin films that has continuous and regular overlapping of the laser spots forms the sharp edge scribe. In addition there were no collars and ridges left at the scribe edge. The scribed line width, which is important parameter in the patterning of the

to Contents, page 1

semiconductor single or multi layers, is mainly depending on laser power and pulse frequency. Laser power varies in this study is from 20 W to 2 W and the pulse frequency varies from 1 Hz to 80 Hz. Fig. 3 shows the obtained comparative results of the scribed line of Mo thin film (60 nm) by varying the laser power and pulse frequency. By increasing the laser power and laser pulse frequency, the scribed line width increases. For higher laser power i.e. 20 W, the width of the scribed line was 240 μm at 1 Hz pulse frequency and 390 μm at 80 Hz pulse frequency. While at lower laser power i.e. 2 W, the increase in the width of the scribed line was negligible.

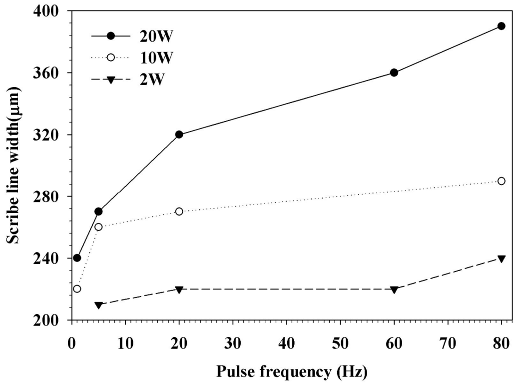


Fig. 3 *– The variation of the scribed line width of Mo thin film (60 nm) as a function of laser pulse frequency shows that as the laser power and the pulse frequency were set at minimum level, the scribe line width is lower*

The scribed line width should be kept minimum for its future use in the patterning the different semiconductor devices. From Fig. 3, we said that at 2 W laser power and 1 Hz pulse frequency we get the minimum scribe width i.e. 210 μm. The 210 μm is still a higher value. Therefore, by using the optimal laser parameters for 60 nm Mo film thickness i.e. 2 W laser power and 1 Hz pulse frequency we scribed the higher thickness of Mo thin film.

Fig. 4 shows the variation in the scribe line width as a function of pulse frequency. In this scribing, we found that as the pulse frequency reduces the scribe line width reduces. The minimum scribing width was 30 μm for 2 W laser power and 1 Hz pulse frequency. We had also scribe the different thickness of Mo thin films at 2 W laser power and 1 Hz pulse frequency and got the similar kind of variation. From that variation, we can say that as the thickness of the Mo thin films increases the scribed line width decreases, too, this is shown in Fig. 5.

At low power (2 W) and low frequency (1 Hz) we have achieve the minimum scribe line width i.e. 40 μm (Fig. 5a). In order to achieve lowest scribe line width without any presence of walls and collars in the scribed area, we scribe the higher thickness Mo films at 0.2 w laser power and 1 Hz pulse frequency, as shown in Fig. 5b. At 0.2 w laser power and 1 Hz pulse frequency the minimum scribe line width is 38 mm. We also try the lower thickness of the Mo thin film for the 0.2 w laser power, but due to the lower thickness, the films experienced a stress, so, the scribe line was distorted or not uniform. Such scribe lines observed using the polarization microscope is shown in Fig. 6.

to Contents, page 1

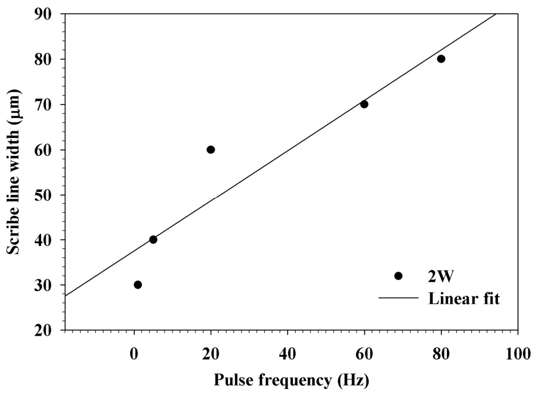


***Fig. 4** – The variation of the scribed line width of Mo thin film (800 nm) as a function of laser pulse frequency shows that as the pulse frequency was reduces; the scribe line width is decreases*

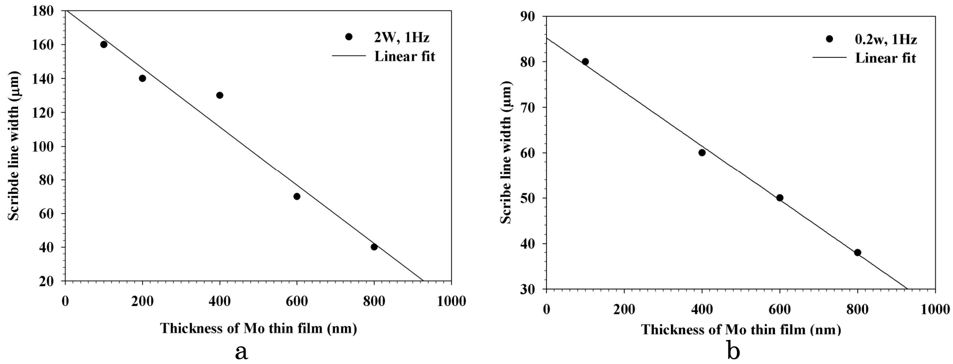


***Fig. 5** – Scribed line thickness is decreases as the thickness of the Mo thin film increases at 2 W laser power and 1 Hz pulse frequency (a) while at the same frequency and 0.2 w laser power the minimum scribe line width observed (b)*

3.2 Scribe quality

However, the great results were extremely rare and seemed to occur within a process window of miniscule proportions. During the extensive experimentation only a handful of 40 μm scribes were achieved with good results. Nonetheless, the excellent quality of these scribes invite to speculation as to whether some alterations could be made to the system in order to achieve such scribes. Scribe results often terrible, with big shards and cracks. The high energy scribe results that are referred to as good on the other hand are not square. Viewed from the top they have a typical pulse-to-pulse appearance of slightly overlapping circular holes. If the work piece speed is too high the pulses will loose their overlap and bridges will occur. If the work piece slows down the “lips” that protrude into the scribe tend to flake up more readily.

to Contents, page 1

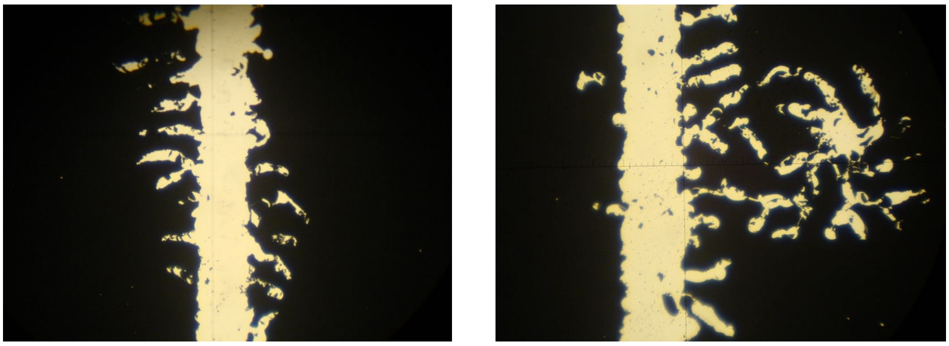


Fig. 6 *– Stress introduce in the lower thickness of Mo thin films while scribing so, the scribe line becomes non-uniform*

In order to achieve the better quality of the scribe line we vary the laser power and pulse frequency for the different thickness of the Mo thin films. Ablation with a train of laser pulses per spot defines the scribe quality. Theoretically, the pulse train of the laser scribe the material is shown in Fig. 7. Some calculations were made in order to establish the geometrical situation [8, 9] during the laser pulse.

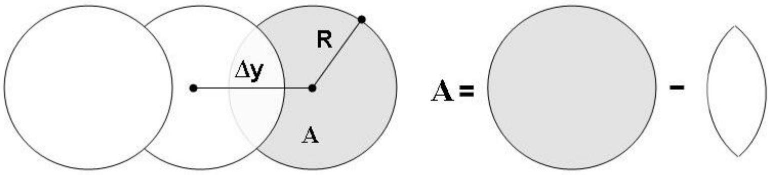


Fig. 7 *– Geometrical representation of exposed area*

Given that the laser spot has a radius, R, and is repeated with a frequency onto a sample that is moving with the speed, υ, one can express the surface area A of each consecutive pulse as the spot size minus the overlap of pulses (see Fig. 7). This area can be considered as the area of film removed per pulse.

$$A = \pi R^2 - 2\left(R^2 \arccos\left(\frac{\Delta y}{2R}\right) - \frac{\Delta y}{2}\sqrt{R^2 - \frac{\Delta y^2}{4}} \right) \tag{1}$$

where Δy is the distance between consecutive pulse centers.

$$\Delta y = \upsilon \times \frac{1}{f} \tag{2}$$

The obtained exposed area (scribe line width) of the scribe line of different thickness of the Mo thin films is shown in Fig. 8.

to Contents, page 1

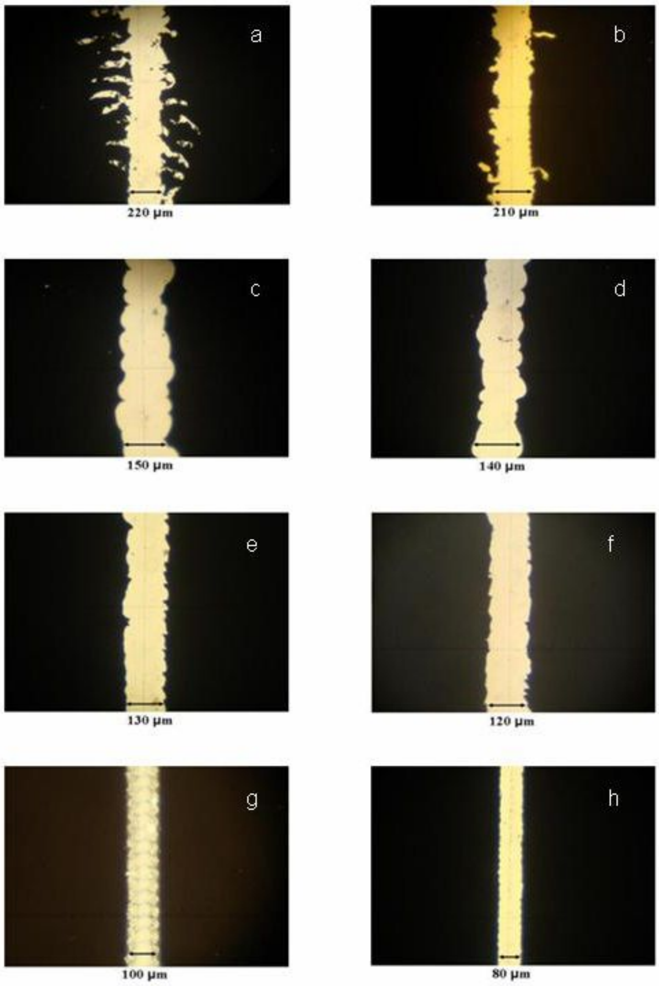


***Fig. 8** – Scribe line width of Mo thin films of different thickness i.e. from (a) to (h) is* 60 *nm to* 800 *nm, results of an average laser power is 2 W and 5 Hz laser pulse frequency*

As increasing the thickness of the Mo thin films, we saw from the Fig. 8, that the scribe line width decreases. At the higher thickness of the Mo thin films, we got the better scribe area as well as the minimum side edges. For 800 nm thin Mo films, the minimum scribe line width i.e. 80 μm was obtain (Fig. 8h). The maximum scribe line width, i.e. 220 μm, with poor scribe area is observed at 60 nm thin Mo films (Fig. 8a).

4. CONCLUSIONS

The presented work shows a process of laser scribing for Mo thin-film of different thickness (60 nm to 800 nm) has been considered. Laser parameters have been determined that provide reproducible, good scribes, that do not present any unwanted bridges. Scribing is performed at energy

to Contents, page 1

densities of around 1 mJ at 20 kHz, with a pulse-to-pulse overlap by the relation between laser power and laser pulse frequency. Considering the laser scribing step, molybdenum having a higher thickness yield better results compare to the lower thickness. Lower thickness of Mo shows the non uniform distribution or dull scribe edges of the scribe line. The scribe line width is observed minimum (80 μm) at higher thickness (800 nm) and maximum (390 μm) at lower thickness (60 nm).

The authors are thankful to UGC-DRS (file no. 530/2/DRS/2007(SAP-1)) for providing an electrical measurement facility.

REFERENCES

1. A.D. Compaan, I. Matulionis, and S. Nakade, *Final Technical Progress Report,* NREL/SR-520-24842 (1998).
2. K.L. Chopra, *Thin Film Phenomenon* (McGraw-hill: 1969).
3. C. Solanki, Solar Photovoltaics: Fundamentals, Technologies and Applications (PHI Learning Pvt. Ltd: 2009).
4. K.L. Chopra, I. Kaur, *Thin Film Device Applications* (NewYork: Plenum Press: 1983).
5. V. Ramanathan, L.A. Russell, C.H. Liu, and P.V. Meyers, *Solar Cells* **28**, 129 (1990).
6. J.G. O'Dowd, N.R. Johnson, R.S. Oswald, and F. Willing, *Proceeding 23rd IEEE PVSC*, 926 (1993).
7. O.S. Heavens, *Thin Film Physics* (Edinburgh: T. & A. Constable Ltd: 1970).
8. R. Arya, *Solarex Final Technical Report. NREL*, contract ZN-I-19019-4 (1995).
9. W.M. Steen, J. Mazumder, *Laser Material Processing*, (Springer-Verlag London Limited: 2010).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.17-25

PACS numbers: 61.05.cp, 81.05.Ea

THE STUDY OF MICROSTRUCTURE OF III-V POLAR ON NON-POLAR HETEROSTRUCTURES BY HRXRD

Ravi Kumar**, Tapas Ganguli, Vijay Chouhan, V.K. Dixit

Raja Ramanna Centre for Advanced Technology,
452013, Indore, India
* E-mail: ravi@rrcat.gov.in

*In this article, we report on the detailed high resolution x-ray diffraction data analysis of three GaAs films deposited by metal organic vapour phase epitaxy on Si substrates. In the GaAs/Si films the effect of anti phase domains is seen by the selective broadening of (*002*) and (*006*) reflections. Further as the (*006*) reflection is a very weak reflection, such films cannot be analyzed by conventional Williamson-Hall plots using (*002*), (*004*) and (*006*) reflections. We find that using (*111*), (*333*) and (*444*) reflections it is possible to use the standard Williamson-Hall analysis and extract parameters related to the microstructure of the films. We have also carried out the analysis to determine the tilt and twist between the mosaic blocks after correcting for the effects of the finite lateral coherence length.*

Keywords: *HRXRD, GaAs/Si, ANTIPHASE DOMAINS, MICROSTRUCTURE, MISMATCHED EPITAXY.*

(Received 04 February 2011, in final form 21 March 2011)

1. INTRODUCTION

High Resolution X-Ray Diffraction (HRXRD) is a powerful non-destructive tool for the analysis of the microstructure of epitaxial systems. There have been several works in recent times, where HRXRD has been used for the understanding of the microstructure of heteroepitaxial layers, like in GaN [1-3], InN [4, 5], ZnO [6], etc.

Williamson Hall (W-H) plots have been used for the analysis of both epitaxial layers and polycrystalline systems for a long time [7]. There has been a large body of work where W-H plots have been used for the analysis of the Lateral Coherence Length (LCL), tilt, vertical coherence lengths and microstrain in epitaxial and oriented polycrystalline films with wurtzite structure, e.g. epitaxial GaN and AlGaN [3], epitaxial InN [4, 5], polycrystalline GaN [8] and polycrystalline ZnO [9]. In all these analysis, the authors have used the reflections from planes perpendicular to the growth direction namely, (0002) (0004) and (0006). There are very few reports on the application W-H plots for the analysis of the microstructure of epitaxial zinc-blende structures deposited on mismatched substrates. In one of the detailed reports, Neumann et al. have used the W-H plots to analyze GaAs/Si system [10]. This system has another complication which is related to the growth of the polar material (namely GaAs) on non-polar Si substrates resulting on the generation of Anti phase domains (APDs). The presence of APDs has been seen in several systems including metallic alloys and is known to broaden selective reflections [10, 11]. In the case of the zinc-

to Contents, page 1

blende systems, the presence of APDs result in the broadening of reflections like (002) and (006) where the expression for the scattering amplitude has the difference of the contributions from the group III and group V atoms [10]. Additionally the (006) reflection is a weak reflection in zinc-blende systems. Further in the case of lattice mismatched zinc-blende epitaxial layers like GaP/GaAs [20], GaSb/GaAs [16], InAs/GaAs etc. [17] which are of contemporary research interest for optoelectronic devices and where APDs are not an issue, the (006) reflection is a very weak, thereby resulting in errors in the determination of the widths. Thus, the W-H analysis using the (002), (004) and (006) reflections cannot be used. In this work, we show that a set of planes parallel to the (111) plane can be alternatively used for the W-H analysis, thereby enabling us the study the microstructure of the zinc-blende layers with very weak (006) reflections.

Following the work of Srikant et al. [12], there has been a lot of work related to the analysis of tilt and twist between the mosaic blocks in lattice mismatched systems for wurtzite systems [13]. However, to the best of our knowledge, there has been no attempt to analyze the tilt and twist between the mosaic blocks in zinc-blende systems. We have carried out a detailed analysis of the tilt and twist using the method proposed by Lee et al. [3] after carrying out LCL corrections.

Although we have looked only at samples of GaAs/Si, the analysis procedure is very general and can be used for any zinc-blende system with APDs like GaP/Si [14,19], GaSb/Si [15] or without APDs like GaSb/GaAs [16]. Thus the proposed methods have a broad range of applicability.

2. EXPERIMENTAL AND ANALYSIS TECHNIQUES

HRXRD measurements have been carried out on a Panalytical X'Pert PRO MRD system. The measurements were made in either the skew-symmetric or the symmetric geometry for different planes (Planes of the type $(00l)$ were analyzed in the symmetric geometry and of the type (h, k, l) with h and/or $k \neq 0$, were analyzed in the skew-symmetric geometry). A hybrid monochromator (Goebel's mirror with a four-bounce crystal monochromator), which gives $CuK\alpha_1$ (wavelength = 1.54056 Å) output with a beam divergence of ~ 20 arcsecs, was used for making the measurements. For GaAs/Si samples a three-bounce collimator (also referred to as triple-axis attachment) is placed in front of the detector to ensure an acceptance angle of ~ 12 arcsecs. The ω - scans in both the symmetric and skew-symmetric geometries have been recorded in the triple-axis geometry for the GaAs/Si samples.

The recorded data is converted in q - space for further analysis using the following relations:

$$q_x = \left(\frac{1}{\lambda}\right)\left[\cos\omega - \cos\left(2\theta - \omega\right)\right] \quad (1)$$

$$q_z = \left(\frac{1}{\lambda}\right)\left[\sin\omega + \sin\left(2\theta - \omega\right)\right] \quad (2)$$

to Contents, page 1

where, λ is the wavelength of the incident X-ray, ω is the angle the sample makes with the sample surface and the 2θ is angle of deviation of the diffracted beam from the incident beam direction.

The ω - scan widths are generally represented by Gaussian line shape profile with two main broadening mechanisms namely: finite LCLs and angle misorientation between the mosaic blocks. The LCL is an average of a large distribution of random sizes of the mosaic blocks. Similarly the angle between the mosaic blocks due to tilts and twists are also from a random distribution of dislocations thereby resulting in a Gaussian broadening. In such a case the total width of the ω - scans can be written as:

$$\left(\Delta q_{obs}\left(\omega\right)\right)^2 = \left(\Delta q_{(00l)LCL}\right)^2 + \left(\alpha_{tilt} \cdot q\right)^2 \quad (3)$$

where, $\Delta q_{(00l)LCL}$ is the broadening due to the finite LCL and α_{tilt}*q is the broadening due to the finite angle between the mosaic blocks. In our case where the instrumental broadening is of the order of 20 arc secs and the measured widths are of the order of 300 arc secs, the instrumental broadening effects have been neglected. However, this is strictly not the case in all the samples. The ω - scans have a pseudo-Voigt line profile that is given by the equation:

$$y = y_o + A\left[f\frac{\left(2w/\pi\right)}{4\left(x-x_c\right)^2 + w^2} + \left(1-f\right)\frac{\sqrt{4\ln 2}}{\sqrt{\pi}w}e^{-\frac{4\ln 2}{w^2}\left(x-x_c\right)^2}\right] \quad (4)$$

where, y_o is a constant, A is the amplitude, f is the fraction of Lorentzian component in the pseudo-Voigt profile ($0 < f < 1$), ω is the width of the curve, and x_c is the peak position. A perfect Lorentzian curve is given by $f = 1$ and a perfect Gaussian is given by $f = 0$.

The addition of two pseudo-Voigt profiles with widths W_1 and W_2 gives a pseudo-Voigt profile with width W, given by the relation:

$$W^n = W_1^n + W_2^n \quad (5)$$

where, $n = 1 + (1 - f)^2$ [12]. For a pseudo-Voigt profile, $1 < n < 2$.

According to the conventional analysis of the ω - scans by W-H plots (Conventionally, the above method has been applied for a set of planes parallel to the sample surface, namely (002), (004) and (006)), the contributions due to the finite lateral size and the tilt are added. The tilt contribution is q dependant (proportional to q) and the lateral size contribution is independent of q. The addition of these two factors for the case of a pseudo-Voigt profile can be written as:

$$\left(\Delta q_{obs}\left(\omega\right)\right)^n = \left(\Delta q_{(00l)LCL}\right)^n + \left(\alpha_{tilt} * q\right)^n \quad (6)$$

where, $\Delta q_{obs}(\omega)$ is the total broadening of the ω - scan peak in the q - space, $\Delta q_{(00l)LCL}$ is the contribution of the broadening due to finite LCL and α_{tilt} is the tilt between the mosaic blocks in the film.

to Contents, page 1

This analysis is however expected to fail in the presence of APDs in the films. The presence of APDs broaden the selective reflections namely (002) and (006) thereby making the equation 6 unusable. Further, in zinc-blende system, (006) is a very weak reflection, which makes the determination of the width highly erroneous.

To avoid this problem, a different set of parallel planes whose reciprocal lattice vector is inclined by some angle with the surface normal, have been used for the analysis. The (111), (333) and (444) reflections, which make an angle (ψ) of 54.73° with the plane parallel to the surface, have been used. It is important that none of these planes are affected by APDs and all of them are allowed reflections that make the data quite strong and the FWHM can thus be determined without significant errors. The ω - scans for these planes have been recorded in the skew-symmetry geometry to obtain the LCL and the angular broadening in the epilayers. The results related to this above modified analysis of the W-H plots are discussed in the next section.

A very important set of parameter that determines the quality of mosaic layer is the tilt and twist between the mosaic blocks of the epitaxial layer. Tilt and twist are out-of-plane and in-plane misorientation of the mosaic blocks in the epitaxial layer. To obtain the values of tilt and twist, ω - scans are recorded for the planes whose angle of inclination with the substrate normal varies from 0° to 90° [12]. The width of x-ray peak obtained by ω - scan for plane for which $\psi = 0°$ is tilt value and width for $\psi = 90°$ is the twist value. As the case of $\psi = 90°$ cannot be recorded in the reflecting geometry, an extrapolation scheme was first used by Srikant et al. [12]. However, the effect of finite LCL is not considered, as all the films analyzed in Ref. [12], had relatively large LCLs. For the set of (1,1,1) planes, the ω - scan broadening may be written as:

$$\left(\Delta q_{obs}\left(\omega\right)\right)^{n} = \left(\Delta q_{(lll)LCL}\right)^{n} + \left(\alpha\left(l,l,l\right) * q\right)^{n} \tag{7}$$

where, $\alpha(1,1,1)$ is the broadening of the (1,1,1) reflection due to a combination of tilt and twist. $\Delta q_{(lll)LCL}$ is the y-axis intercept of the W-H plot for ω - scans using (111), (333) and (444) reflections. The contribution of LCL to be peak broadening is eliminated from the ω - scan of all the skew-symmetric reflections. Thus the LCL corrected angular broadening α (h, k, l), for any (h, k, l) reflection in the skew-symmetric geometry may be written as:

$$\alpha\left(h,k,l\right) = \left(\left(\frac{\Delta q_{obs}}{q}\right)^{n} - \left(\frac{\Delta q_{(lll)LCL}}{q}\right)^{n}\right)^{1/n} \tag{8}$$

After the correction for the LCL, the expression for the width as a function of ψ may be written as [3,18]:

$$\alpha\left(h,k,l\right)^{n} = \left(\alpha_{tilt}\cos\psi\right)^{n} + \left(\alpha_{twist}\sin\psi\right)^{n} \tag{9}$$

The GaAs layers (samples “a”, “b” and “c”) were deposited by Metal Organic Vapour Phase Epitaxy (MOVPE) (AIXTRON AIX200) on (001)

to Contents, page 1

oriented Si substrates. The samples "a", "b" and "c", have been deposited using a two step growth technique with different V/III ratios. The buffer layers were grown at 450 °C and at 400 °C and the epilayers at 650 °C, 670 °C, and 690 °C respectively for samples "a", "b" and "c". The thickness of the samples were determined by a thickness profilometer model Alpha-step IQ (KLA Tencor make). We present a detailed HRXRD study for the above three epilayers. The analysis is very general and can be used for any III-V epilayer especially those that are highly lattice mismatched where the issues of tilt, twist and LCL are very important.

3. RESULTS AND DISCUSSION

Fig. 1 shows the intensity vs. q_x curves derived from ω - scans using eqn. 1 and 2 for **(002)**, **(004)** and **(006)** reflection of layer "a". The pseudo-Voigt fitting of the curves using eqn. 4 are shown by the overlaying lines in the figure. The weak reflection **(006)** reflection is very noisy and the fitting is erroneous. The observation is similar for all the remaining samples.

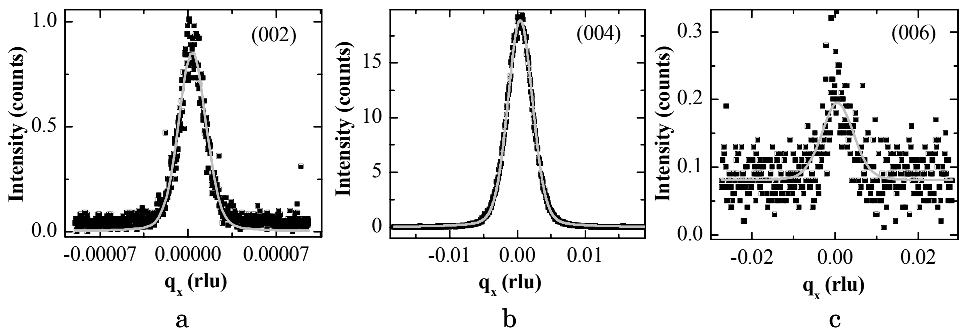


Fig. 1 *– The intensity vs. q_x curves for (002), (004) and (006) reflection of sample "a". The fitting of the curve is shown by overlaying line*

Fig. 2a, b, and c shows the W-H plots using the ω - scans for the layers "a", "b" and "c" respectively using **(002)**, **(004)** and **(006)** reflections. The value of "n" used in the plots are determined from the pseudo-Voigt fitting of the ω - scan profiles in the q - space using equations 1, 2 and 4. One finds that for all the layers, straight line fits, as expected from W-H plots are not obtained (Fig. 2a-c). This observation is attributed to two reasons. First, the values of the width of the **(006)** reflection, which is a forbidden reflection for zinc-blende structure, may be erroneous due to its very small intensity. Secondly, it may be noted that **(002)** and **(006)** reflections, are selectively broadened by the presence of APDs in the GaAs/Si samples i.e. epilayers "a", "b" and "c", thereby making the straight line fitting of the W-H plots impossible.

We have also made W-H plots using the ω - scans of the **(111)**, **(333)** and **(444)** reflections in the skew-symmetric geometry, from the epilayers "a", "b" and "c" in Fig. 3a-c respectively. We find that straight-line fits are obtained in all the cases thereby enabling the possibly of using this analysis for this set of reflections. The value of the LCL and α **(1, 1, 1)** for the films determined from the straight line fitting (also shown in Fig. 3a-c) is given

to Contents, page 1

in Table 1. Thus we find that the choice of the diffraction planes in this case is extremely important and LCL values can only be determined in these films using the set of (1, 1, 1) planes. $\alpha(1, 1, 1)$ values determined from these plots are a combination of tilt and twist in the layers and thus need to be analyzed further based on the procedure mentioned in the previous section.

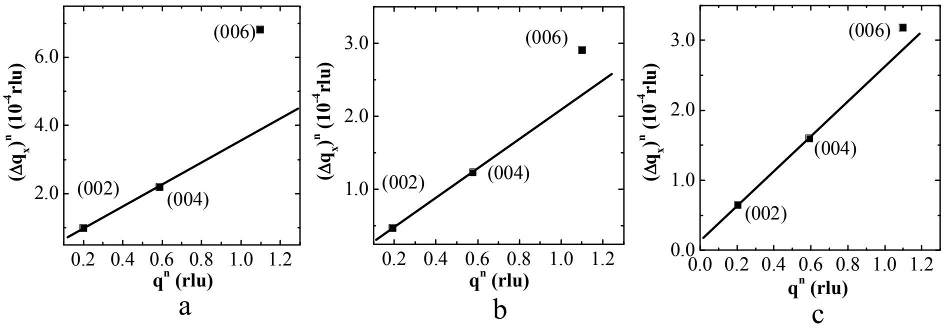


***Fig. 2** – W-H plots for samples "a", "b" and "c" using the (002), (004) and (006) reflections for the ω - scan widths. The straight line is used to guide the reader's eye*

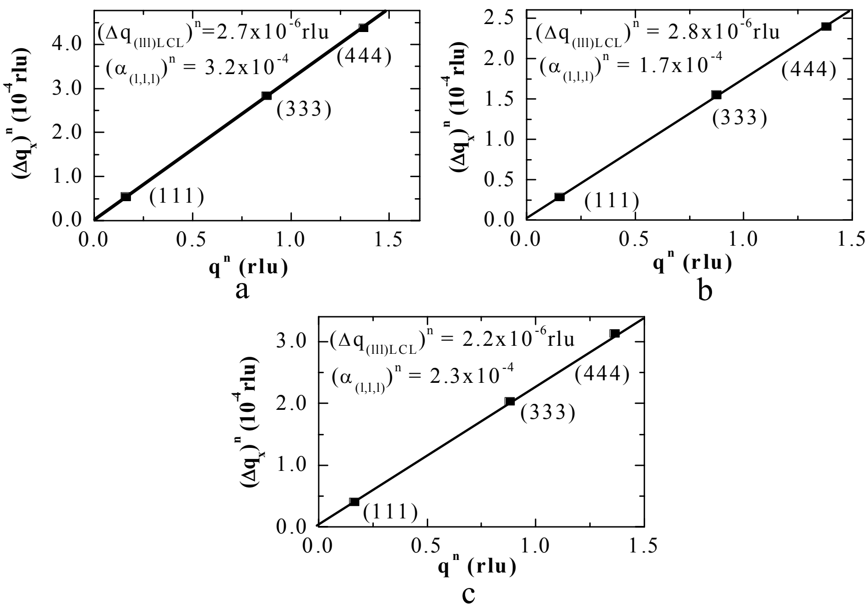


***Fig. 3** – W-H plots for samples "a", "b" and "c" using the (111), (333) and (444) reflections for the ω - scan widths. The straight line shows the linear fitting of the points*

We estimate the tilt (out-of-plane misorientation of the blocks) and the twist (in-plane misorientation of the blocks) by making a series of reflections in the skew-symmetric geometry and plotting the width of these reflections as a function of ψ. The y axis (ω - scan width) shown in Fig. 4, is equivalent to $\Delta q_{obs}/q$ of eqn. 8. This is strictly true only for small ω widths, which is

to Contents, page 1

almost always the case with epitaxial films. These results are shown in Fig. 4a-c for the epilayes “a”, “b” and ‘c” respectively. The widths without the finite length corrections are shown in the inset of each figure. It is clear that the correction due to the finite length effects is mandatory and the data after the correction due to the finite lengths is seen to fit a monotonic line as expected from Eqn. 9. The fits to Eqn. 9 for all the films are shown in Fig 4a-c. The values of the tilt and twist are shown in the Table 1.

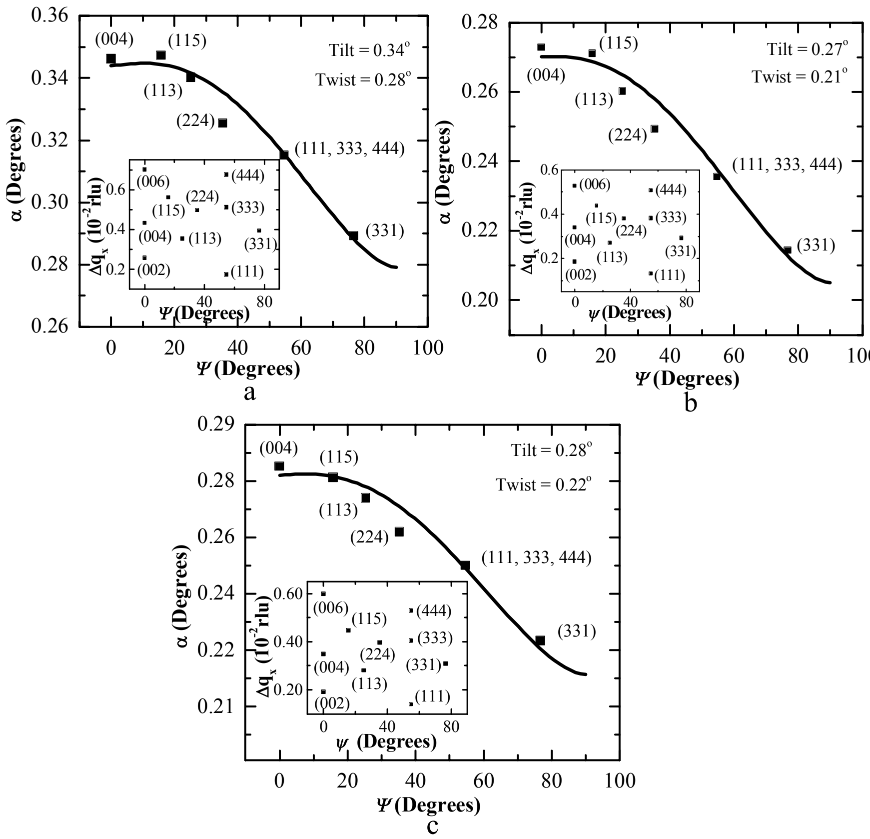


Fig. 4 *– Widths of ω - scan data recorded in the skew-symmetric geometry corrected for the finite coherence lengths for the planes shown as a function of ψ. The data for the samples “a”, “b” and “c” are shown. Inset shows widths of ω - scan data recorded in the skew-symmetric geometry for the planes shown in the inset as a function of ψ. The data for the samples “a”, “b” and “c” are shown*

The tilt and twist values obtained for the films are directly related to the dislocation density of the films. This method has been the standard method for the determination of the dislocation density for III-Nitride epilayers where the systems are in almost all the cases grown on highly mismatched substrates. However for the case of III-V systems, almost all the work has been carried out on matched or nearly matched systems and thus there issues of tilt and twist between the mosaic blocks do not arise. However there are a few systems where this issue is important like the ones studied

to Contents, page 1

in this work namely: GaAs/Si. The most common type of dislocation present in these systems are the ones that have a Burger's vector $b = a/2$ (0, 1, 1) and line of dislocation along: (1, − 1, 0) direction, commonly referred to as the 60° dislocation. These are mixed dislocations that give rise to both tilt and twist in the epilayers. The details of the effects of the dislocations, their contribution to the tilt and twist and their relation to the measured TEM data will be a subject of a future work.

Table 1 *– The results obtained for the various epilayers "a", "b and "c"*

Sample Name	Sample Thickness (μm)	LCL determined From W-H plots (μm)	Tilt Determined from W-H plots (Degrees)	Twist Determined from W-H plots (Degrees)
a	0.31	0.40	0.34	0.28
b	0.31	0.31	0.28	0.22
c	0.33	0.51	0.27	0.21

4. CONCLUSION

Standard W-H Analysis using (001) reflections (l = 2, 4, 6) cannot be used with ease for the zinc-blende systems due to APDs and/or weak reflections. A modified W-H Analysis is presented in this work that can be used for a zinc-blende system. This approach is specifically applicable for the cases, where the films have APDs and/or some of the reflections in the W-H plot are not allowed and hence very weak. The LCL and the angular broadening of the mosaic blocks are determined from the W-H plots. The tilt and twist between the mosaic blocks in the epilayers is also obtained after correcting for the LCL.

The authors thank Dr. T.K. Sharma and Mr. S.D. Singh for help in the planning and growth of the samples, Dr. M.P. Joshi and Dr. Rajmohan for the thickness measurements and Dr. S.M. Oak for helpful discussions, encouragement and support. The authors also thank Mr. U.K. Ghosh and Mr. A. Khakha for their help during the growth of the samples.

REFERENCES

1. R. Chierchia, T. Buttcher, H. Heinke, S. Einfeldt, S. Figge, D. Hommel, *J. Appl. Phys.* **93,** 8918 (2003).
2. M.E Vickers, M.J Kappers, R Datta, C McAleese, T.M Smeeton, F.D.G Rayment, C.J Humphreys, *J. Phys. D: Appl. Phys.* **38,** A99 (2005).
3. S.R. Lee, A.M. West, A.A. Allerman, K.E. Waldrip, D.M. Follstaedt, P.P. Proventio, D.D. Koleske, C.R. Abernathy, *Appl. Phys. Lett.* **86,** 241904 (2005).
4. T. Ganguli, A. Kadir, M. Gokhale, R. Kumar, A.P. Shah, B.M. Arora, A. Bhattacharya, *J. Cryst. Growth* **310,** 4942 (2008).
5. X.L. Zhu, L.W. Guo, N.S. Yu, J.F. Yan, M.Z. Peng, J. Zhang, H.Q. Jia, H. Chen, J.M. Zhou, *J. Cryst. Growth* **306,** 292 (2007).
6. S. Singh, R. Kumar, T. Ganguli, R.S. Srinivasa, S.S. Major, *J. Cryst. Growth* **310,** 4640 (2008).
7. G.K. Williamson, W.H. Hall, *Acta Metall.* **1,** 22 (1953).

to Contents, page 1

8. B.S. Yadav, S. Singh, T. Ganguli, R Kumar, S.S. Major, R.S. Srinivasa, *Thin Solid Films* **517,** 488 (2008).
9. S. Singh, T. Ganguli, R. Kumar, R.S. Srinivasa, S.S. Major, *Thin Solid Films* **517,** 661 (2008).
10. D.A. Neumann, H. Zabel, R. Fischer, H. Morkoз, *J. Appl. Phys.* **61,** 1023 (1987).
11. B.E. Warren, *X-ray diffraction* (New York, Courier Dover Publications: 1990).
12. V. Srikant, J.S. Speck D.R. Clarke, *J. Appl. Phys.* **82,** 4286 (1997).
13. T. Metzger, R. Hцpler, E. Born, O. Ambacher, M. Stutzmann, R. Stommer, M. Schuster, H. Gobel, S. Christiansen, M. Albrecht, H.P. Strunk, *Philos. Mag. A* **77,** 1013 (1998).
14. V.K. Dixit, T. Ganguli, T.K. Sharma, S.D. Singh, R. Kumar, S. Porwal, P. Tiwari, A. Ingale, S.M. Oak, *J. Cryst. Growth* **310,** 3428 (2008).
15. S.H. Vajargah, M Couillard, Y. Shao, S. Tavakoli, R. Kleiman, J. Preston, G.A. Botton, *Microsc. Microanal.* **16,** 1338 (2010).
16. K. Li, J. Lin, A.T.S. Wee, K.L. Tan, Z.C. Feng, J.B. Webb, *Appl. Surf. Sci.* **99,** 59 (1996).
17. H. Yamaguchi, K. Kanisawa, S. Miyashita, Y. Hirayama, *Physica E* **23,** 285 (2004).
18. M.A. Moram, M.E. Vickers, *Rep. Prog. Phys.* **72,** 036502 (2009).
19. V.K. Dixit, T. Ganguli, T.K. Sharma, R. Kumar, S. Porwal, V. Shukla, A. Ingale, P. Tiwari, A.K. Nath, *J. Cryst. Growth* **293,** 5 (2006).
20. Z. Zhaochun, Q. Xiaoyan, C. Deliang, K. Xianggui, H. Baibiao, J. Minhua, *Mater. Sci. Eng. B* **77,** 24 (2000).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.26-40

PACS numbers: 68.55.Ln, 81.15.Cd

ENHANCEMENT OF NMP DEGRADATION UNDER UV LIGHT BY NITROGEN-DOPED TiO_2 THIN FILMS USING A DESIGN OF EXPERIMENT

***H. Fakhouri*[1], *W. Smith*[1], *J. Pulpytel*[1], *A. Zolfaghari*[2], *H. Mortaheb*[2], *F. Meshkini*[2], *R. Jafari*[2], *F. Arefi-Khonsari*[1]**

[1] Laboratoire de Genie des Procedes Plasmas et Traitements de Surfaces (LGPPTS), Universite Pierre et Marie Curie (UPMC),
ENSCP, 11, Rue Pierre et Marie Curie, 75231, Paris, Cedex 05, France
E-mail: jerome-pulpytel@enscp.fr

[2] Chemistry and Chemical Engineering Research Center of Iran,
Tehran, Iran

Doping nitrogen within TiO_2 is an effective way to enhance visible light photocatalysis due to a direct electron excitation from the N_{2p} states within the band gap. However, nitrogen doping is not always efficient for UV photocatalytic activity. Here, different structures of N-doped TiO_2 (TiO_xN_y) have been prepared by reactive RF (13.56 MHz) magnetron sputtering. The morphological, optical, structural, and photocatalytic properties of the films have been studied in order to investigate the competitive effect of the morphology and the chemical composition on the efficiency of the photocatalytic activity. The variation of surface wettability of the film over time in the dark and under visible and UV irradiation was also studied. The reduction in wettability by dark storage can be explained by the adsorption of hydrocarbon contamination on the thin film's surface. Additionally, from water contact angle experiments, it was found that these films developed hydrophilic properties upon UV and visible illumination. The photoinduced change in the contact angle of water was due to the removal of hydrocarbon contamination on the surface and also the photo-oxidation of the water droplet. Samples prepared at high pressure gave the best photocatalytic activity, even though the deposition rate was lower at higher pressures (lower film thicknesses), due to the high specific surface area and the optimal presence of TiO_xN_y crystals in the lattice. However, at low pressure, the TiN crystals became more predominant, and acted as recombination centers for the photo-generated charge carriers. A design of experiments was used in order to optimize the deposition parameters to have the best photocatalytic activity. The high photocatalytic activity under UV light was found to be due to the introduction of discrete energy levels within the band gap, the increased sample wettability, and the higher specific surface area. However, the post annealing process did not effect the activity under UV irradiation. Using the response surface methodology, RSM, based on a design of experiment, DOE, we are able to achieve a good understanding of the complex processes involved in the deposition of the thin films and their effect on the photocatalytic activity.

Keywords: *TIO_2 THIN FILM, DOE, RSM (RESPONSE SURFACE METHOD), NMP, PHOTOCATALYSIS, NITROGEN DOPING, SPUTTERING.*

(Received 04 February 2011)

to Contents, page 1

1. INTRODUCTION

TiO_2 has shown tremendous applications as an effective photocatalyst under UV light irradiation [1-7]. It is a readily available material that is low cost, has good thermal and aqueous stability and resistivity, and is adverse to corrosion and photocorrosion. TiO_2 has shown the ability to photodegrade many organic pollutants dissolved in water via UV irradiation. Among organic pollutants, N-methyl-2-pyrrolidone, (NMP), is one of the most common found in the wastewater of many factories due to its common and cast usage. NMP is an important, versatile solvent and reaction medium for the chemical industry such as petrochemical, plastics, coating, agricultural, and electronics, because of its low volatility, thermal stability, high polarity and aprotic, noncorrosive properties [8]. However, due to its large band gap TiO_2 absorbs visible light poorly, and thus has been an ineffective visible-light photocatalyst. Many attempts have been made to improve the photocatalytic activity of TiO_2 under UV and/or visible light irradiation. In particular, the focus has been on decreasing the band gap energy of TiO_2 (about 3.2 eV), such that it is suitable to absorb visible light, i.e. 1.6 - 3.0 eV, particularly at 2.5 eV, which corresponds to the solar maximum output. The efforts to modify the optical properties of TiO_2 have been attempted by three primary methods; (1) doping TiO_2 with transition metal ions (Cr, V, Mn, etc.) in order to produce intermediate states in the band-gap close to the conduction band [9-13], or/and with anionic dopants (F,C, N, ...) close to the valence band [14-17]. (2) sensitization by attaching photo-sensitizers such as organic substances, semiconductors, and metal halides that are able to absorb visible light [18-20], and (3) the formation of reduced TiO_2 containing oxygen vacancies [21-23], including reduction with hydrogen [24]. While there are a plethora of available doping methods, some processes can hurt the overall photoactivity of the TiO_2 samples. For instance, some doping methods suffer from thermal or photoactivite instability, an increase in the recombination between the photogenerated electron-hole pairs, and the poor reproducibility of the photocatalytic activity [25-26].

In particular, Asahi et al. [27] initially set three requirements to achieve visible-light photoactivity for TiO_2: (i) doping should produce states within the band gap of TiO_2 that absorb visible light; (ii) the conduction band CB minimum and the dopant states must be above the H_2/H_2O reduction level to ensure photoreductive activity; and (iii) sufficient overlap between the intergap and band states of the photocatalyst in order to provide a fast transfer of photoexcited carriers to the reactive states at the surface within their lifetime. Metal dopants do not meet conditions (ii) and (iii) as they produce localized *d*-states deep within the band gap of TiO_2 and tend to act more as electron-hole pair recombination centers. In practice, metal doping has shown both positive and negative effects on the photocatalytic activity of TiO_2, under both UV and visible light [28-29].

An important factor in the applicable photoactivity of doped-TiO_2 is based on the density of states (DOS) that make up the photoactive energy levels, i.e. the valence band, the conduction band, and the doped energy band level. The density of states determine the transition probability for photoexcitation from either the valence band or dopant energy level, to the conduction band level, which is governed by Fermi's Golden Rule. DOS calculations have been carried out using anionic dopants such as C, N, F, P [30] and S [31] (however, Sulfur doping is not commonly employed due to its large ionic radius even if it can narrow the band gap [31]). The most interesting results are that with carbon

to Contents, page 1

[15] and nitrogen, which can generate intermediate impurity states above the maximum of the valance band within the band gap [16, 30, 32-34]. Interstitial and substitutional carbon doping may decrease the TiO_2 band gap to 2.32 eV which is even less than the known Rutile band gap (about 3.0 eV), so the photoinduced charge carriers can be generated more easily under visible light irradiation in C-doped TiO_2 [15]. However, the mobility of the photogenerated holes may be reduced, which will not improve the photooxidation activity which takes place via a cationic exchange, even if the measured photocurrent is increased anionically [28].

Nitrogen doped TiO_2 has received significant attention, and the results have yielded many contradictions. The DOS results (Asahi et al. for example) suggest that the substitutional nitrogen doping is more effective than interstitial doping (and the mixture of both substitution and interstitial) due to the introduction of localized N_{2p} states just above the valance band which can mix with O_{2p} states (N—O bonding), allowing the absorption of visible light up to 500 nm. Sangwook Lee et al [35] found that interstitial N doping creates defects at levels deep within the band gap which disturb the charge transfer in the TiO_2 nanoparticles. However, other works found that interstitial N-doping is more effective than substitutional doping [35-36]. Diwald et al. [37-38] found that interstitial doping shows more photoactivity due to the suppressing photo-threshold in opposition of substitutional dopants. Other results have shown that a multi-type N doping (interstitial *and* substitutional) gives very good photocatalytic activity in visible light [39].

Although some of the past literature has found discrepancies in the nature and function of nitrogen doping, there are a few common themes that have become evident. While doping TiO_2 with nitrogen can decrease the band gap, and thus allow an improvement in visible light absorption, the nitrogen band energies that allow this absorption also can act as recombination centers that decrease the lifetime of photogenerated charge carriers, which decrease the overall photocatalytic efficiency. There exists a delicate balance in the concentration (and nature) of N-doping in TiO_2 that must be achieved in order to maximize the visible light absorption, while minimizing the electron-hole recombination. The purpose of this study is to find the sputtering deposition parameters that allow for the optimal optical and photocatlytic properties of N-doped TiO_2 films.

Sputtering deposition is one of the most promising techniques for uniform coatings over a large area as it has the advantage of producing a well controlled film in terms of morphology, crystallinity and stochiometry. The reactive RF magnetron sputtering of TiO_2 is of great importance because it has shown the ability to prevent target poisoning [40-45], and it also permits desirable and reproducible properties of the deposited films with good deposition rate at different substrate temperatures, as well as a good homogeneity of thickness and composition of films on a large scale with a flat surface. In addition, reactive RF magnetron sputtering of titanium dioxide has a high degree of flexibility to control the structural and physical properties of the films by varying several process parameters such as the sputtering pressure, sputtering power, target to substrate distance, reactive gas partial pressure, substrate temperature and using DC and RF substrate biasing [42, 44, 46-49].

The study of the individual processing parameters affecting photocatalytic activity is quite complicated due to the abundance of physical, chemical, and electronic procedures that take place, as well as the abundance of sputtering

to Contents, page 1

controls, detailed above. The use of response surface methodology, RSM, based on design of experiment is one way to simplify complex processes since it allows multiple physical characteristics to be compared simultaneously. The resolution of the empirical polynomial equation that determines the RSM can be used to understand and predict the optimization of the photocatalytic activity of N-doped TiO_2 by comparing the material properties with the process parameters and functional activity. In this work, we used a design of experiment to investigate the morphological, structural, and optical properties of N-doped TiO_2 thin films with various nitrogen concentrations, and how they affect the UV light photocatalysis of NMP.

2. EXPERIMENTAL

2.1 Thin films preparation

The un-doped and N-doped TiO_2 thin films were prepared by a RF reactive magnetron sputtering system. A metallic Ti target (50 mm diameter) having a purity of 99.95%, was sputtered in a reactive gas atmosphere containing Ar, O_2 and N_2. Different nitrogen to oxygen ratios were used, ranging from 0 to 14 %, while the argon flow rate was fixed to 30 sccm. The substrates used were electro-polished stainless steel, Si (100) wafers and ordinary microscope glass slides. Prior to deposition, the non metallic substrates were ultrasonically cleaned with acetone, ethyl alcohol and then de-ionized water for 20 minutes each. The RF cathode power was 200 W for all the depositions in order to minimize the target poisoning effect [41-42] since titania is a semi-conducting oxide, and can easily contaminate the target. Using optical emission spectroscopy (OES), we followed the Ti emission lines at 517 and 521 nm, and controlled the TiO_2 deposition parameters in order to have a fixed intensity of the Ti peaks during the deposition period. We stabilized the deposition rate in the intermediate region between two stable sputtering modes, reactive and metallic. The substrate holder was maintained at 300°C. The distance between the substrate holder and the target was fixed at 100 mm. The initial base pressure was between 3 to 7×10^{-6} torr, while the total working pressure was between 3 to 14 mtorr. Prior to each deposition, the target was pre-sputtered in Argon atmosphere for 10 minutes to remove any contaminations from the target surface. The film thicknesses were measured, and ranged from 80 to 500 nm. Some of the films were annealed for 1 hour at 450 °C in air atmosphere. The TiO_xN_y films in this paper were labeled as following: Pr for pressure, and N for the ratio of the nitrogen to oxygen flow rates.

2.2 Contact angle measurement

Contact angle measurements were performed by measuring the angles at which the liquid/vapor interface from a 6 μl water droplet meets the liquid/solid sample surface. Initial tests were performed for each sample under no light irradiation and the contact angle variation was studied over the course of two months. Further tests were conducted for different intervals of illumination with the full spectra of a Halogen and UV lamps. The time needed for a drop of water to return to the initial contact angle (of the fresh sample) after 2 months of dark storage was measured. A set of three 6 μl droplets were used for each surface, and the average contact angle over time was measured and averaged.

to Contents, page 1

2.3 Film structure and morphology

The crystal structure of thin films was characterized by X-ray diffraction (XRD) (X'Pert Pro PW3040-Pro, Panalytical Inc.) using Cu K-R radiation in Bragg-Brentano (θ-2θ) configuration. The average crystallite size was determined by means of standard θ-2θ XRD scans using the Scherrer Formula. X'Pert High Score pattern processing was used to collect and process the data. The Raman spectra were recorded in a backscattering geometry using a Renishaw InVia Raman-microscope system and the 633 nm excitation wavelength of a red laser, focused on a spot size of the order of 3 μm. Film morphology and thickness were measured by SEM-FEG.

2.4 Photocatalytic activity

N-methyl-2-pyrrolidone (NMP) was oxidized in the presence of the prepared thin films using an 11W lamp (Philips TUV low pressure Hg lamp for near monochromatic output at 254 nm) which was placed at the center of the reactor shielded by a vertically immersed quartz jacket. All the experiments were performed at constant temperature. NMP was obtained from Merck with purity of > 99 %. The concentration of NMP and produced intermediates were monitored during the experiments using HPLC-UV (KNAUER, Chromgate 3.1) with isocratic elution (3 - 10 % acetonitrile, 97 - 90 % water, 1 ml/min) on a reversed-phase Nucleosil C_{18} column. The detection wavelength was 214 nm.

2.5 Design of experiment

In order to determine how the individual process and material parameters affect the photocatalytic activity of TiO_2 thin films, we used the Doehlert design with 2 factors (nitrogen flow rate and sputtering pressure). In order to achieve the desired set of data, we used 7 experiments focused around a central parameter value, including one in the center [51-52]. We estimated the optimized conditions by the Response Surface Methodology, RSM, based on a second order model.

3. RESULTS AND DISCUSSION

3.1 Contact angle measurement

Depending on the deposition parameters, the prepared samples displayed two types of contact angle shapes (within 2 months). Fig. 1a and b present an example of the contact angle variation related to two types of samples prepared at high pressure and low pressure. It is apparent that the contact angle increases with time with two principal slopes before reaching a stationary value. Fig. 1b shows the contact angle variation after different intervals of illumination with the full spectra of a Halogen and UV lamp up to return to the initial value of the contact angle (of the fresh sample).

We distinguished two groups (types) of contact angle trends according to the controlled pressure during the film deposition. Here we present only an example of samples prepared at high and low pressure respectively (Fig. 1a). At high pressure (more than 9 mtorr) the contact angle increases slowly and steadily up to a stable value about 10 % more than the initial contact angle value. However, samples prepared at low pressure show a different variation

to Contents, page 1

following two mean slopes, a first order and a second order one up to a stable value up to 8 times over the initial contact angle value. Under visible irradiation, the wettability increases suddenly after a specific irradiation time which increases with increasing pressure up to reaching to the initial contact angle. The same conclusion can be seen under UV irradiation but the wettability is increased within a shorter time of UV irradiation and the surfaces become completely super hydrophilic regardless of the deposition pressure.

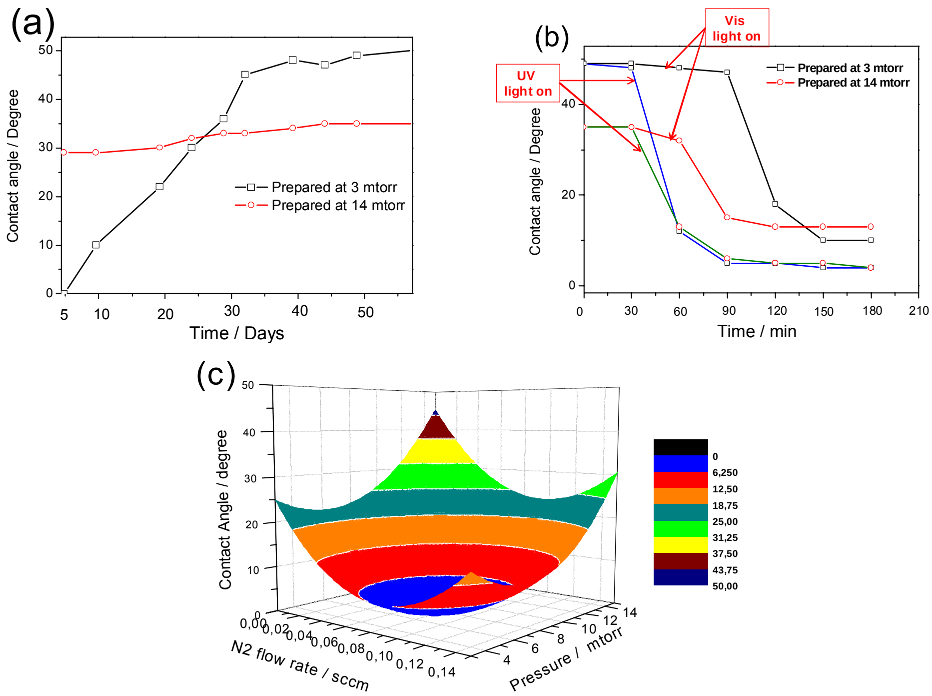


***Fig. 1** – (a) The contact angle evolution during 2 months, being stored in the dark, (b) the change in contact angle under visible and UV light irradiation, and (c) the measured contact angle for different deposition pressures and nitrogen flow rates (the oxygen flow rate is fixed at 1.5 sccm, and contact angle is measured after 5 days of deposition)*

The contact angle increased after storing the sample in the dark, possibly due to the adsorption of hydrocarbons on the photocatalyst surface. As the illumination exposure time is increased, the contact angle decreases from an initial value of ~50° (in the dark) to a final value of ~10є after two hours of illumination (the illumination period was in the absence of water droplet, and we used 6 μl de-ionized water droplet for each contact angle measurement on the same sample). These results are in agreement with literature [53], especially the model proposed by Wang et al. [50], which reported the photo induced change in contact angle of water to be due to the adsorption of hydrocarbon contamination on the surface of the thin film.

The contact angle enhancement, under irradiation, is due to two phenomena. First, under irradiation, the degradation of adsorbed hydrocarbons took place on the doped titania surface through the generation of OH radicals. Second,

to Contents, page 1

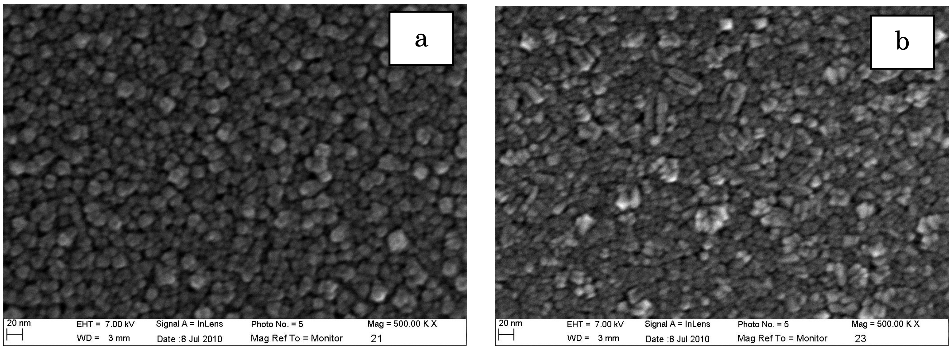
illumination initiates the photo-oxidation of the water droplet, leading to a larger wettability of the surface. We used the wettability data, measured one week after the samples deposition, to modulate the response with two parameters, the pressure and the nitrogen flow rate. Fig. 1c presents the 3-D empirical iso-response curves, to identify the experimental region for minimum contact angle.

3.2 Impact of the sputtering pressure and film morphology

The microstructure of TiO_2 layers has a pronounced effect on the photocatalytic activity. The small pore size of the thin film allows only the top surface to be photocatalytically active. The sputtering process is able to grow column-like grain structures, and the spacing between columns can be controlled by the sputtering pressure [54]. The increased porosity increases the effective surface and active sites. For example, we can calculate the lateral ($S.A._{LAT}$) to geometrical ($S.A._{GEO}$) surface ratio for the film deposited at 14 mtorr (Fig. 2a) which has the minimum thickness (80 nm),

$$\frac{S.A._{LAT}}{S.A._{GEO}} = \frac{4\pi Dd}{\pi D^2} \tag{1}$$

where (d) is the film thickness and (D) is the column diameter. Using equation (1) and the morophological parameters measured by SEM, it is



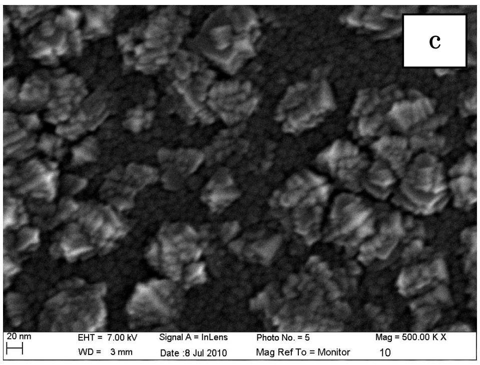


***Fig. 2** – SEM-FEG images of 3 thin films, N0.04 Pr14 (a) N0.04 Pr 4 (b) and N0.15 Pr4 (c)*

apparant that the available surface area is enhanced 10-fold when increasing the pressure from 4 to 14 mtorr, although it is important to note that the ratio is proportional to the film thickness.

to Contents, page 1

As the TiO_2 crystals coalesce before reaching the substrate surface, their shape and size will stabilize. Normally, the crystal size increases up to a threshold that will further inhibit its growth. This maximum grain size is important for the inherent surface area of the thin film (the higher grain size would generate a higher porosity and a higher specific area) . The larger grain size would enhance the crystallinity which increases the adsorption of the pollutants, and thus show improved photodegradation. From Fig. 2 it is clear that at higher deposition pressure (comparing (a) with (b)), the film is more porous having a higher effective surface area. At a higher nitrogen flow rate (higher nitrogen concentration) it is possible to see the nucleation of TiO_xN_y clusters and this is more evident at high working pressure.

Fig. 3 shows the RMS for thickness in function of pressure and nitrogen flow rate. The thickness (and also the deposition rate) decreases sharply with the sputtering pressure. However, it is apparent that the nitrogen does not have any important impact on the deposition rate.

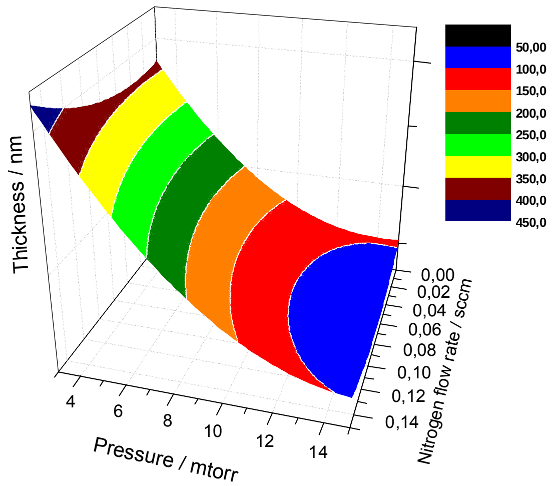


Fig. 3 *– Thickness of N doped TiO_2 thin films and iso-reponse curves (based on pressure and Nitrogen flow rate) from different samples data. Oxygen flow rate is fixed to be 1.5 sccm*

3.3 XRD analysis and Raman spectroscopy

The photocatalytic activity of TiO_2 depends on a variety of parameters including, but not limited to, the TiO_2 crystalline structure, crystal/particle size, surface morphology and the concentration/nature of the doping element. The crystalline phase of TiO_2 remains one of the principal factors that determines the photocatalytic performance. The anatase form of TiO_2 is preferable over rutile since anatase is found to be highly active in the photo-oxidation of organics when oxygen is used as an acceptor of photogenerated electrons [7].

X-ray diffraction scans (Fig. 4) provide evidence that anatase is the primary crystalline phase that develops in the thin films. By changing the level of atomic nitrogen doping (resulting from different N_2 flows, so a different N_2:O_2 reactive mixture ratio is achieved) during the sputtering process, there is a slight decrease in the anatase phase and increase in the development of the TiO_xN_y phase. By increasing the sputtering pressure, the

to Contents, page 1

anatase peaks, especially (101) and (004), located at $2\theta = 25.4°$ and $2\theta = 38.0°$ respectively, become less intense and broader. This indicates that the crystal size and the refractive index are decreasing [54]. At the same time, the (103), (112), and (105) orientations of the anatase crystal, located at $2\theta = 37.1°$, $38.7°$ and $54.1°$ respectively, disappeared as the higher pressure depositions were used. The (105) anatase peak was increased by increasing the nitrogen incorporation in the lattice. It is not surprising that the anatase (101) decreases rapidly by increasing the nitrogen flow rate. Many authors reported a crystal modification when substituting nitrogen into the anatase TiO_2 crystal. The wide peak between $2\theta = 41°$ and $44°$ is attributed to the presence of the TiO_xN_y phase accompanied with TiN_x and/or TiO_{55}. By changing the nitrogen flow rate, we can identify a remarkable formation of TiN crystals accompanied with oxygen vacancies in the lattice for nitrogen flow rates between 0.07 and 0.11 sccm, and it is more distinguishable at a higher sputtering pressure. Oxygen is a much more reactive element than nitrogen [45] so altering the oxygen to nitrogen flow rate's ratio can put considerable strains on the TiN crystallization process, explaining the observed broad diffraction peak at higher oxygen concentrations.

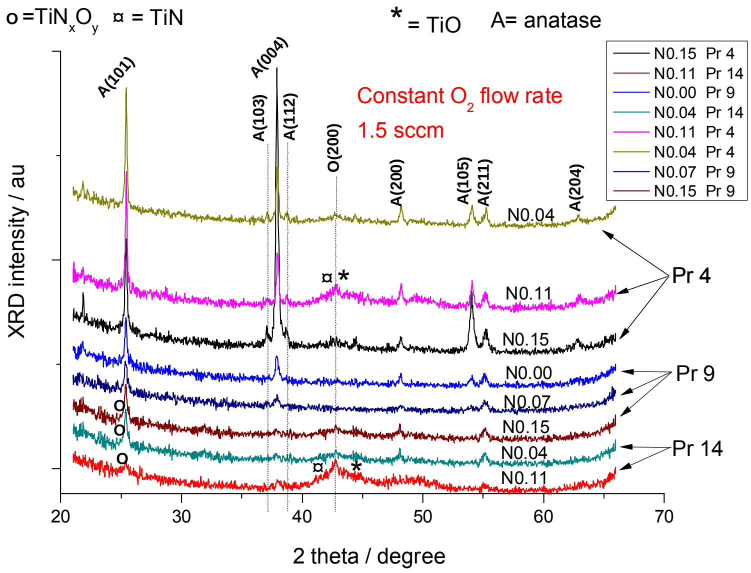


***Fig. 4** – XRD data obtained from TiO_2 films prepared at different pressures (Pr) and different Nitrogen flow rates* (0.04 *to* 0.15 *sccm)*

The avereage crystal size in each film was calculated from the anatase (101) peak using the Scherrer formula,

$$d = K\lambda/(\beta\cos\theta) \tag{2}$$

where d is the crystal size in nm, K is a constant, λ is the XRD wavelength (in our case 1.54 Å), β is the full width at half maximum (FWHM), and θ is the diffraction angle. To try and understand the competitive nature of

to Contents, page 1

nitrogen incorporation and the sputtering pressure, we plotted the surface response of the crystal size versus both variables (Fig. 5). It is apparent that the crystal size decreases by increasing the nitrogen flow rate, especially at higher pressure. However, this variation is less remarkable when working at low pressure (less than 9 mTorr).

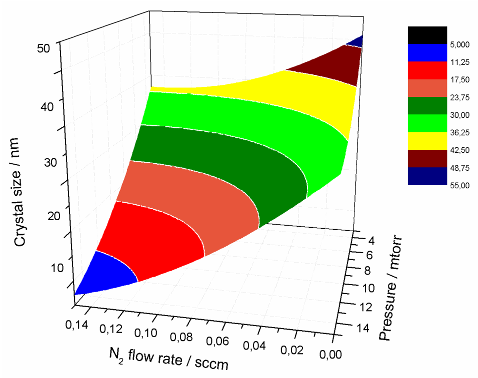


Fig. 5 *– Surface response of the crystal size versus deposition pressure and nitrogen flow rate*

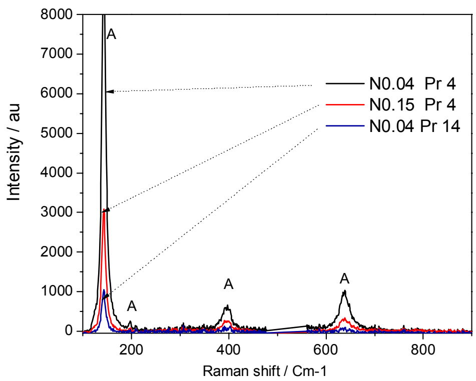


Fig. 5 *– Raman spectra of a selection of deposited films on silicon wafer*

The Raman shift spectra of the different thin films were measured and are shown in Fig. 6. All the samples show the typical Raman shift spectra of anatase TiO_2 [56-57], comprised of the peaks at 144 cm^{-1} (E_g, O–Ti–O bending mode), 198 cm^{-1} (B_{1g}/A_{1g}), 398 cm^{-1} (B_{1g}), and 640 cm^{-1} (E_g, Ti–O bond stretching mode). In addition to the optical band modes found by this method, no phonon line shifts were observed . Raman spectroscopic analysis can provide important information on the Ti–O bond lengths of the N-doped TiO_2 crystals from the stretching wave numbers which is related to the force constants and lengths of the bonding [2, 35]. There is no red shift in the E_g peak (at 640 cm^{-1} corresponding to the stretching vibrational mode of the Ti–O bond) when comparing the pure TiO_2 sample to any of the doped samples. The absence of the red shifting peak means that nitrogen is only substitutionally doped, and not found interstitially.

to Contents, page 1

3.4 Band gap calculation

The optical properties of all samples were studied by measuring the transmission spectra in the UV-Visible range (210 to 1010 nm). The band gap values were calculated by the Tauc's equation using an indirect allowed transitions ($m = 2$),

$$\alpha h\nu = B(h\nu - E_g)^m, \tag{3}$$

where α is the absorption coefficient, B is a constant called the edge width parameter, αh is the photon energy, E_g is the optical band gap, and m is a constant which is determined by the optical transition, either direct (m = S) or indirect (m = 2). The optical band gap is obtained from the extrapolation of the linear plots of $(\alpha h\nu)^{1/2}$ vs $h\nu$, where the x-intercept is the band gap energy. The calculated band gap values were estimated to be between 2.5 and 3.42 eV for the N-doped TiO_2 samples. This large difference is due to the different nitrogen concentrations and nitrogen positions in the lattice depending on the deposition pressure and the morphology of the various thin films. Oxygen vacancies can also have an important role on the band gap red shift.

The RSM of the modulated band gap (from the calculated ones) is presented in Fig. 7. It is clear that the conditions corresponding to the minimum band gap are in the same interval that display higher hydrophilicity. This is a very interesting correlation, suggesting that the position of the doped nitrogen energy levels within the band gap can have a substantial effect on the surface states of the material, which in turn affect its wettability.

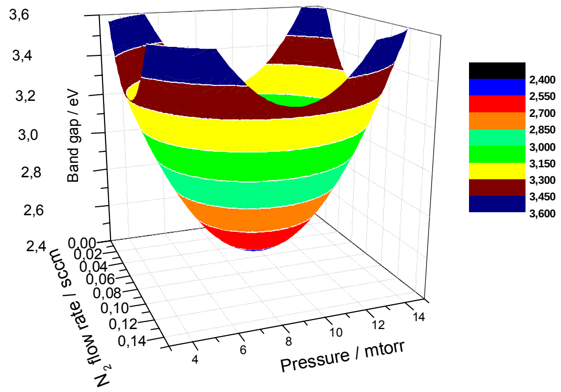


***Fig. 7** – Modelling of the band gap calculated from the UV-Vis absorption spectra of glass coated by N doped TiO_2*

3.5 NMP degradation and optimization of the sputtering process

Using colored dyes as indicators in photocatalytic reactions can be difficult because it is sometime hard to separate the observed activity between the dye's photocatalytic decomposition, and the reduction of the dye due to the electron transfer reactions. In both cases we would get the dye's decoloration as that found with methylene blue [60]. We selected NMP because it is a pollutant in industrial waste water, and its concentration can be easily controlled by HPLC, which allows the observation of sub-components in real

to Contents, page 1

time. Fig. 8 shows the NMP degradation with and without the presence of different thin films photocatalysts under UV irradiation. It is important to note that the pure TiO_2 has an intermediate effect compared to the other doped samples. This means that in some sputtering conditions, N-doped TiO_2 can enhance the photocatalytic activity under UV irradiation compared to pure TiO_2, while other sputtering conditions can have a negative effect on the photocatalytic reactions under UV illumination. This can reasonably explain the contradictions found in literature about the efficiency of nitrogen doping under UV illumination (see introduction).

The interaction of the plasma pressure and the nitrogen doping with respect to the photocatalytic decay of NMP in the TiO_2 thin films is shown in Fig. 9. We can distinguish two directions for the minimum NMP degradation time (optimum conditions). First, at higher pressure and lower nitrogen incorporation, porous thinner films are deposited with high active surface sites and show good efficiency for the photocatalysis in UV irradiation. On the other hand, a lower deposition pressure and higher nitrogen doping gives thicker films with less surface active sites, but still shows good photocatalytic activity. At an intermediate nitrogen doping and sputtering pressure, the deposition mechanism is more complicated.

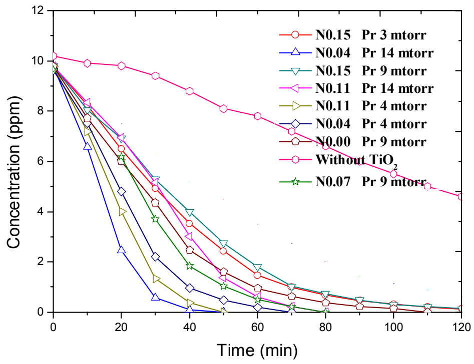


Fig. 8 – NMP degradation time with different deposited TiO_2 samples

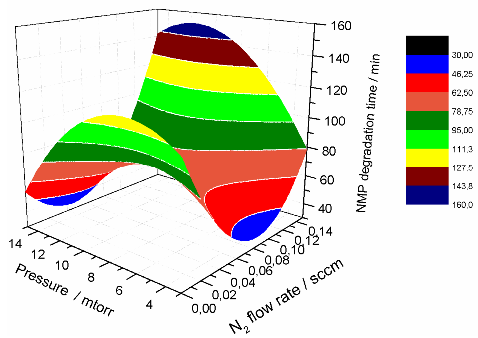


Fig. 9 – Modelisation of NMP degradation time using the measured time up to total degradation by different samples

It is well known that TiN and oxygen vacancies can increase the conductivity, and can also act as recombination centers in TiO_2 [61, 62]. At an

to Contents, page 1

intermediate nitrogen flow rate, we found a remarkable formation of TiN crystals in the lattice, as shown in the XRD pattern in Fig. 4. This can explain the decreased photocatalytic efficiency at this interval of nitrogen flow rates (0.07 to 0.11 sccm), as TiN is not a good photocatalyst because it does not have semiconductor electronic properties, or good absorption in the UV range.

4. CONCLUSION

The efficiency of N doping in TiO_2 thin films to improve photocatalytic performance under UV irradiation may have a contradictory effect, which is clear from the literature survey. Using RF reactive sputtering as a deposition technique, we achieved different structures of N-doped TiO_2 (TiO_xN_y). There is a competitive effect of the morphology and the chemical composition on the efficiency of the photocatalytic activity under UV irradiation, which can be seen clearly using the response surface methodology, based on a design of experiment. The variation of surface wettability of the films over time in the dark is due to the hydrocarbon's adsorption on the surface of the TiO_2 thin films. Samples prepared at high pressure can have an optimized photocatalytic activity, even with lower film thicknesses, due to the high specific surface area and the optimal presence of TiO_xN_y crystals in the lattice. This study has shown that there is a delicate balance between the concentration and nature of nitrogen doping in TiO_2 that can lead to both improved and deteriorated photocatalytic activity. However, we have shown that there are optimal deposition conditions to maximize the photocatalytic efficiency of N-doped TiO_2, which can serve as a solid background for further investigations.

Support for this project from the Syrian Higher Institute for Applied Sciences and Technology and from the Gundishapur (EGIDE) are gratefully acknowledged.

REFERENCE

1. U.G. Akpan, B.H. Hameed, *J. Hazard. Mater.* **170,** 520 (2009).
2. M. Kitano, M. Matsuoka, M. Ueshima, M. Anpo, *Appl. Catal. A: Gen.* **325,** 1 (2007).
3. F. Han, V.S. Rao Kambala, M. Srinivasan, *Appl. Catal. A: Gen.* **359,** 25 (2009).
4. I.N. Martyanov, K.J. Klabunde, *J. Catal.* **225,** 408 (2004).
5. H. Szymanowski, A. Sobczyk-Guzenda, A. Rylski, W. Jakubowski, M. Gazicki-Lipman, U. Herbert , F. Olcaytug, *Thin Solid Films*, **515,** 5275 (2007).
6. E. Aubry, M.N. Ghazzal, V. Demange, N. Chaoui, D. Robert, A. Billard, *Surf. Coat. Tech.* **201,** 7706 (2007).
7. A. Fujishima, X. Zhang, D.A. Tryk, *Surf. Sci. Rep.* **63,** 515 (2008).
8. B. Elvers, S. Hawkins, W. Russey, G. Schulz, *Encyclopedia of industrial chemistry* (VCH: New York: 1993).
9. D. Kaczmarek, D. Wojcieszak, J. Domaradzki, E. Prociow, *J. Phys.: Conf. Ser.* **146,** 012019 (2009).
10. Z. Liu, Y. Zhou, Z. Li, Yichao Wang, C. Ge, *J. Univ. Sci. Technol. B* **14,** 552 (2007).
11. Y. Li, M. Ma, X. Wang, Z. Li, *Surf. Coat. Tech.* **204,** 1353 (2010).
12. I.M. Arabatzis, T. Stergiopoulosa, M.C. Bernardc, D. Laboud, S.G. Neophytidesd, P. Falaras *Appl. Catal. B: Env.* **42,** 187 (2003).
13. T.C. Kaspar, T. Droubay, V. Shutthanandan, S.M. Heald, C.M. Wang, D.E. McCready, S. Thevuthasan, *Phys. Rev. B* **73,** 155327 (2006).
14. T. Yamaki, T. Sumita, S. Yamamoto, *J. Mater. Sci. Lett.* **21,** 33 (2002).

to Contents, page 1

15. S.U.M Khan, M. Al-Shahry, W.B. Ingler, *Science* **297,** 2243 (2002).
16. A.V. Emeline, V.N. Kuznetsov, V.K. Rybchuk, N. Serpone, *Int. J. Photoenergy,* 258394 (2008).
17. K. Obata, H. Irie, K. Hashimoto, *Chem. Phys.* **339,** 124 (2007).
18. W. Smith, Y.P. Zhao, *J. Phys. Chem. C* 112, 19635 (2008).
19. C. Hu, X. Hu, L. Wang, J. Qu, A. Wang, *Environ. Sci. Technol.* **40,** 7903 (2006).
20. P.V. Kamat, *Chem. Rev.* **93,** 267 (1993).
21. I. Nakamura, N. Negishi, S. Kutsuna, T. Ihara, S. Sugihara, K. Takeuchi, *J. Mol. Catal. A: Chem.* **161,** 205 (2000).
22. P. Zeman, S. Takabayashi, *Surf. Coat. Tech.* **153,** 93 (2002).
23. T. Ihara, M. Miyoshi, Y. Iriyama, O. Matsumoto and S. Sugihara, *Appl. Catal. B: Env.* **42,** 403 (2003).
24. Y. Liu , X. Chen, J. Li, C. Burda, *Chemosphere* **61,** 11 (2005).
25. E. Borgarello, J. Kiwi, M. Gratzel, M. Pelizzetti, M. Visca, *J. Am. Chem. Soc.* **104,** 2996 (1982).
26. L. Spanhel, H. Weller, A. Henglein, *J. Am. Chem. Soc.* **109,** 3183 (1987).
27. R. Asahi, T. Morikawa, T. Ohwaki, K. Aoki, Y. Taga, *Science* **293,** 269 (2001).
28. T.L. Thompson, J.T. Yates, *Chem. Rev.* **106,** 4428 (2006).
29. W. Smith, S. Mao, G. Lu, A. Catlett, J. Chen, Y. Zhao, *Chem. Phys. Lett.* **485,** 171 (2010).
30. T. Morikawa, R. Asahi, T. Ohwaki, K. Aoki, K. Suzuki, Y. Taga, *R&D Review of Toyota CRDL,* **40** No.3 45 (2005).
31. T. Umebayashi, T. Yamaki, H. Itoh, K. Asai, *Appl. Phys. Lett.,* **81,** 454 (2002).
32. C. Di Valentin, E. Finazzi, G. Pacchioni, A. Selloni, S. Livraghi, M.C. Paganini, E. Giamello, *Chem. Phys.* **339,** 44 (2007).
33. D.J. Mowbray, J.I. Martinez, J.M. Garcia Lastra, K.S. Thygesen, K.W. Jacobsen *J. Phys. Chem. C* **113,** 12301 (2009).
34. J. Wang, D.N. Tafen, J.P. Lewis, Z. Hong, A. Manivannan, M. Zhi, M. Li, N. Wu, *J. Am. Chem. Soc.* **131,** 12290 (2009).
35. S. Lee, I. Cho, D.K. Lee, D.W. Kim, T.H. Noh, C.H. Kwak, S. Park, K.S. Honga, J.K. Lee, H.S. Jung, *J. Photoch. Photobio. A* **213,** 129 (2010).
36. S.H. Lee, E. Yamasue, K.N. Ishihara, H. Okumura, *Appl. Catal. B: Env.* **93,** 217 (2010).
37. O. Diwald, T.L. Thompson, T. Zubkov, E.G. Goralski, S.D. Walck, J.T. Yates, *J. Phys. Chem. B* **108,** 6004 (2004).
38. O. Diwald, T.L. Thompson, E.G. Goralski, S.D. Walck, J.T. Yates, *J. Phys. Chem. B* **108,** 52 (2004).
39. F. Dong,W. Zhao, Z. Wu, S. Guo, *J. Hazard. Mater.* **162,** 763 (2009).
40. A. Brudnik, M. Bucko, M. Radecka, A. Trenczek-Zajac, K. Zakrzewska, *Vacuum* **82,** 936 (2008).
41. U. Helmersson, M. Lattemann, J. Bohlmark, A.P. Ehiasarian, J.T. Gudmundsson, *Thin Solid Films* **513,** 1 (2006).
42. W.D. Sproul, D.J. Christie, D.C. Carter, *Thin Solid Films* **491,** 1 (2005).
43. F. Perry, A. Billard, C. Frantz, *Surf. Coat. Tech.* **94-95,** 681 (1997).
44. J.W. Bradley, T. Welzel, (Process Diagnostics), S. Konstantinidis, F. Gaboriau, M. Gaillard, M. Hecq, A. Ricard, (Optical Plasma Diagnostics During Reactive Magnetron Sputtering), *Reactive Sputter Deposition* (Springer Berlin Heidelberg New York: 2008).
45. A. Trenczek-Zajac, M. Radecka, K. Zakrzewska, A. Brudnik, E. Kusior, *J. Power Sources* **194,** 93 (2009).
46. J. Musil, P. Baroch, J. Vlcek, K.H. Nam, J.G. Han, *Thin Solid Films* **475,** 208 (2005).
47. B.C. Trasferetti, C.U. Davanzo, R.A. Zoppi, *Phys. Rev. B* **64,** 125404 (2001).
48. M.S. Wong , H.P. Chou, T.S. Yang, *Thin Solid Films* **494,** 244 (2006).

to Contents, page 1

49. W. Zhou, X. Zhong, X. Wu, L. Yuan, Q. Shu,W. Li, Y. Xia, *J. Phys. D: Appl Phys.* **40,** 219 (2007).
50. C. Wang, H. Groenzin, M.J. Shultz, Langmuir **19,** 7330 (2003).
51. S.L.C. Ferreira, W.N.L. dos Santos, C.M. Quintella, B.B. Neto, J.M. Bosque-Sendra, *Talanta* **63,** 1061 (2004).
52. M.A. Bezerra, R.E. Santelli, E.P. Oliveira, L.S. Villar, L.A. Escaleira, *Talanta* **76,** 965 (2008).
53. X.T. Zhao, K. Sakka, N. Kihara, Y. Takada, M. Arita, M. Masuda, *Microelectr. J.* **36,** 549 (2005).
54. K. Eufinger, D. Poelman, H. Poelman, R. De Gryse, G.B. Marin, *J. Phys. D: Appl. Phys.* **40,** 5232 (2007).
55. N. Martin, R. Sanjines, J. Takadoum, F. Levy, *Surf. Coat. Tech.* **142-144,** 615 (2001).
56. M. Ritala, M. Leskela Erja Nykдnen, P. Soininen, L. Niinisto, *Thin Solid Films* **225,** 288 (1993).
57. W.R. Kaiser, A.N. Christensen, G. Mullervogt, *Phys. Rev. B*, **17** 1095 (1978).
58. L. Popovic, D. de Waal, J.C.A. Boeyens, *J. Raman Spectrosc.* **36,** 2 (2005).
59. T. Ohsaka, F. Izumi, Y. Fujiki, *J. Raman Spectrosc.* **7,** 321 (1978).
60. A. Mills, J. Wang, *J. Photoch. Photobio. A* **127,** 123 (1999).
61. H.M. Yates, M.G. Nolana, D.W. Sheela, M.E. Pemble, *J. Photoch. Photobio. A* **179,** 213 (2006).
62. S. Takeda, S. Suzuki, H. Odaka, H. Hosono, *Thin Solid Films* **392,** 338 (2001).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.41-46

PACS numbers: 68.55 aj, 73.90 + F

EFFECT OF SUBSTRATE TEMPERATURE ON STRUCTURAL AND MORPHOLOGICAL PARAMETERS OF ZnTe THIN FILMS

***K.D. Patel*[1], *G.K. Solanki*[1], *C.J. Panchal*[2], *K.S. Hingarajiya*[1], *J.R. Gandhi*[1]**

[1] Department of Physics, Sardar Patel University
Vallabh Vidyanagar, 388 120, Gujarat, India
E-mail: kdptflspu@yahoo.com

[2] Faculty of Technology, Applied Physics Department,
M.S. University of Baroda, Vadodara 390001, Gujarat, India

Vacuum evaporated thin films of Zinc Telluride (ZnTe) of 5000 Å *thickness have been deposited on glass substrates at different substrate temperatures (*303 *K,* 373 *K,* 448 *K). Structural parameters were obtained using XRD analysis. Atomic Force Microscope (AFM) in non-contact mode has been used to study the surface morphological properties of the deposited thin films. The results obtained from structural and surface morphological studies have been correlated and it is found that the films deposited at higher substrate temperatures possess increasingly good crystallinity and smoother surfaces.*

Keywords: *ZNTE THIN FILM, THERMAL EVAPORATION, XRD, AFM, SUBSTRATE TEMPERATURE.*

(Received 04 February 2011, in final form 17 March 2011)

1. INTRODUCTION

Thin films of II-VI compound semiconductors have drawn researcher's attention for more than four decades. ZnTe is expected to be a promising material for a variety of optoelectronic devices, such as pure green light emitting devices, detectors for various optoelectronic instrumentation, etc. because of its direct wide band gap of 2.26 eV [1-3]. It is also used as terahertz detectors [4, 5] and window material for CdTe based solar cells [6]. Many researchers have used various techniques for the fabrication of ZnTe thin films including Metal organic chemical vapor deposition [7], MBE [8], vacuum evaporation [9-11] R.F. Sputtering [12] and Electrodeposition [13-17]. Among these, thermal evaporation technique offers several advantages including simplicity and cost effectiveness for larger area processing. In this paper we report our results of our investigations on structural and morphological properties dependence on substrate temperature in case of ZnTe thin films deposited using thermal evaporation technique.

2. EXPERIMENTAL

Thin films of ZnTe were deposited on an ultrasonically cleaned glass substrates using thermal evaporation technique, under a vacuum of 10^{-6} Torr at different substrate temperatures (303 K, 373 K and 448 K). A polycrystalline ZnTe (99.999 %, Aldrich make) powder was used to deposit 5000 Å thick films at the deposition rate of 5 Å per second. Structural analysis was made using X-ray diffraction technique (XRD) with the help of

to Contents, page 1

CuKα radiation and the surface morphological study was carried out using atomic force microscope in non-contact mode with tungsten carbide tip.

3. RESULTS AND DISCUSSION

3.1 Structural Analysis

The X-ray diffractograms of the thin films deposited at various substrate temperatures are shown in Fig. 1. It is observed that XRD patterns of all three films show a most preferred orientation along (111) plane. Also the most prominent peak is observed nearly at 25.6° for 2θ, which shows that the deposited films possess cubic structure [18]. Structural parameters of the deposited films are given in Table 1. The sharper and intense peaks, in Fig. 1, for films deposited at increased substrate temperatures exhibit an improved crystalline structure of the films. Particle size values also increases with increase in the substrate temperature.

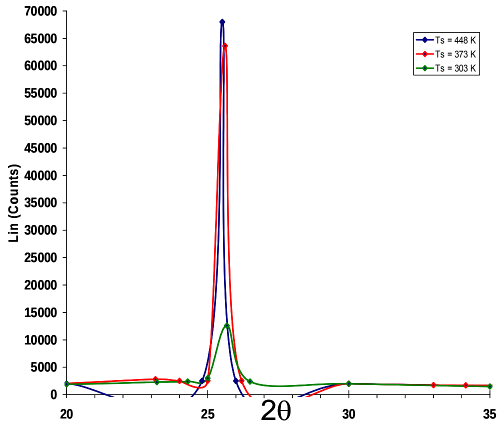


Fig. 1 – *XRD peaks of ZnTe thin films of thickness* 5000 Å, *deposited at various substrate temperatures*

Table 1 – *Structural parameters of ZnTe thin films of thickness* 5000 Å *deposited at various substrate temperatures*

Parameter	T_s = 303 K	T_s = 373 K	T_s = 448 K	JCPDS Values
$a = b = c$ (Å)	6.0038	6.0183	6.0414	6.0700
Unit Cell Volume V (Å)^3	216.4106	217.9844	220.5021	223.6485
X-ray Density ρ (gm × cm^3)	5.9221	5.8794	5.8123	5.7306

3.2 Morphological Analysis

The surface morphology of ZnTe films have been studied using Atomic Force Microscopy. Two dimensional and three dimensional AFM images (5 µm × 5 µm) of the deposited films along with height profile and the power spectrum are shown in Fig. 2.

to Contents, page 1

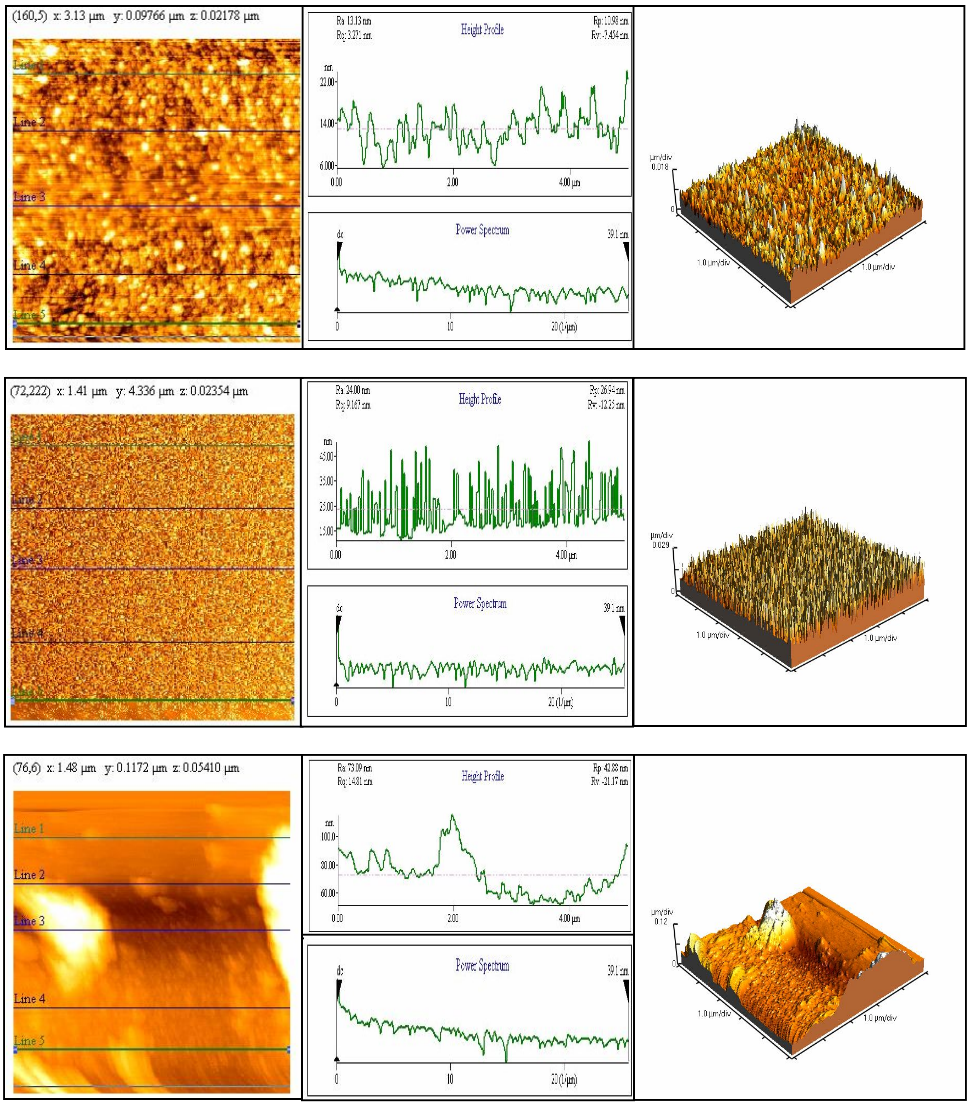
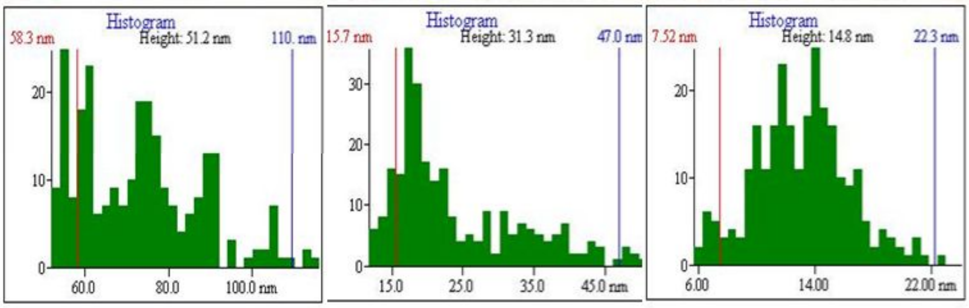


***Fig. 2** – Two and three dimensional images of ZnTe Thin films of Thickness* 5000 Å, *deposited at various substrate temperatures:* $T_s = 303$ *K (a),* $T_s = 373$ *K (b),* $T_s = 448$ *K (c)*

The data obtained from the analysis are given in Table 2. The difference between the rms and average roughness values of line-1 and line-2 (12.87 nm & 24.72 nm) of the film at substrate temperature 303 K is larger than that of the other three lines. This indicates poor smoothness of the film in that area which can be clearly seen in the Fig. 2. In case of the films at substrate temperatures 373 K and 448 K this difference is comparably smaller and is decreases with an increase in substrate temperature, which is the indication of better surface smoothness for the deposited films.

to Contents, page 1



a b c

***Fig. 3** – Histograms of ZnTe thin films of thickness* 5000 Å, *deposited at various substrate temperatures:* T_s = 303 K (a), T_s = 373 K (b), T_s = 448 K (c)

***Table 2** – Data obtained from AFM analysis of ZnTe thin films*

Line	Sub. Tem. (k)	$R_{P\text{-}V}$ (nm)	R_{RMS} (nm)	R_{AVG} (nm)	Mean Ht. (nm)	Med. Ht (nm)	Arc Lth (μm)	Bearing Ratio @30% (nm)	Bearing Ratio @80% (nm)	Peak R_P (nm)	Valley R_V (nm)
01	303	124.7	24.72	12.87	73.15	66.61	5.049	71.21	61.27	109.9	-14.79
	373	49.98	11.52	9.558	23.99	21.48	6.069	28.51	13.59	32.66	-17.31
	448	16.86	3.482	2.865	13.14	12.92	5.032	15.04	10.13	9.970	-6.891
02	303	206.8	53.39	39.21	73.12	47.43	5.076	73.21	37.48	164.4	-42.43
	373	44.03	10.31	8.474	23.99	23.29	6.101	29.23	14.19	26.34	-17.69
	448	20.99	4.201	3.126	13.10	12.07	5.032	14.01	9.77	14.62	-6.366
03	303	210.7	68.35	61.26	73.14	30.45	5.150	113.9	19.04	150.8	-59.92
	373	45.73	10.61	8.846	23.98	22.93	5.997	29.36	13.59	26.73	-19.00
	448	11.16	2.387	1.909	13.12	12.82	5.020	14.27	11.12	6.495	-4.665
04	303	44.15	10.74	9.193	73.11	73.15	5.156	80.12	61.87	223.21	-20.94
	373	48.89	10.27	8.404	24.01	22.44	5.935	28.69	15.04	30.09	-18.79
	448	17.59	3.151	2.540	13.06	13.15	5.039	14.85	10.36	9.676	-7.913
05	303	64.05	14.81	11.97	73.09	72.91	5.150	78.49	59.20	42.88	-21.17
	373	39.18	9.161	7.600	24.00	20.25	5.930	27.71	17.10	26.94	-12.25
	448	18.44	3.271	2.585	13.13	13.20	5.025	14.65	10.41	10.98	-7.454

A difference between peak and valley values ($R_{P\text{-}V}$), obtained from the height profile is listed in Table 2. A large difference in the value of $R_{P\text{-}V}$ for all five lines of the image of film at lower substrate temperature (ranging from 44.15 nm to 210.7 nm) shows a poor surface smoothness of that film. The films deposited at higher substrate temperature show relatively low value of $R_{P\text{-}V}$. Thus it is again confirmed by the difference of peak and valley observations made along five arbitrarily scanned horizontal lines of AFM scan that as substrate temperature increases, the surface smoothness improves.

to Contents, page 1

Bearing ratio is the two dimensional projection of three dimensional surface. It gives a percentage of covered area in a film at the particular height. Thus it shows a length of the particle above a horizontal line throughout the distribution. Bearing ratio allows a comparison of roughness data for all three films and it is listed in Table 2. There is a large difference in these values (ranging from 71.21 nm to 113.9 nm @ 30%) and 19.04 nm to 61.87 nm @ 80 %) for the film deposited at $T_s = 303$ K. For other two films with $T_s = 373$ K and 448 K these differences are comparably smaller.

Power spectrum curve is the important parameter in analysis of any rough surface. It determines the contact area between two solids and can provide both, lateral and longitudinal information. A convenient way to describe surface roughness is to represent it in the term of profile heights $z(x, y)$. For a typical digitized AFM scans, the value of x and y are quantized. Thus the power spectra exhibit the overall surface features of the deposited films as shown in Fig. 2. Looking to all three curves it is clear that the curve for the film at $T_s = 303$ K possess an irregular and spread peaks in comparison to that of the other two films. Peaks in the power spectrum indicate the periodicity of the surface and frequency of each peak gives a length that defines this periodic surface. Spread peak exhibits the deviations from average value. The spectrum for the film at $T_s = 373$ K exhibits a sharp peak showing better surface properties.

Mean height, which is the central value of the roughness profile over the evaluation length, decreases as the substrate temperature of the deposited film is increased, showing a better smoothness of the film surface. The median height which is a mid point on the roughness profile over the evaluation length such that half of the data fall above it and half below it, is also inversely proportional to the substrate temperature of the film. In the obtained data, there is a large variations in median value for the film at $T_s = 303$ K, particularly at line number 2 and 3, which can also be seen from the Fig. 3. The large difference in mean and median values of the film deposited at $T_s = 303$ K shows an asymmetric distribution.

The histogram is a continuous bar diagram in which, each column represents the number of image pixels having the height value in a particular range. The histograms for the deposited films are shown in Fig. 2. It indicates a decrease in height from 51.2 nm to 14.8 nm for the films deposited at $T_s = 303$ K to 448 K.

4. CONCLUSION

ZnTe thin films were successfully deposited on the glass substrates using thermal evaporation method at various substrate temperatures. It is clear from structural data that films show better crystalline structure at the higher substrate temperature. It is also clear from the detailed analysis of various AFM parameters like rms and average roughness, mean and median heights, bearing ratio, peak and valley values, power spectrum density and histograms, that at the higher substrate temperatures the deposited films possesses a better smoothness and crystalline structure on its surface, which supports the XRD analysis.

to Contents, page 1

The authors are thankful to Department of Chemistry, Sardar Patel University, Vallabh Vidyanagar and SICART, Vallabh Vidyanagar for AFM characterization and XRD analysis.

REFERENCES

1. D. Kurbatov, V. Kosyak, M. Kolesnyk, A. Opanasyuk and S. Danilchenko, *Integrated Electronics* **103**, 32 (2008).
2. R. Bhargava, Properties of wide band gap II-VI semiconsuctors- London, United Kingdom: INSPEC, The Institution of Electrical Engineers, 238 (1997).
3. T. Tanaka, K. Saito, M. Nishio, Q. Guo and H. Ogawa, *J. Mater. Sci. Mater. Electron.* **20**, 505 (2008).
4. T. Ishizaki, N. Saito, O. Takai, S. Asakuru, K. Goto and A. Fuwa, *Electrochim. Acta* **50**, 3509 (2005).
5. C. Winnewisser, P.U. Japsen, M. Schall, V. Schiyja. H. Helm, *Appl. Phys. Lett.* **70**, 3069 (1997).
6. P.V.S. Meyers, Proceedings of seventh Commission European Communities Confrrence o Photovolatic Solar Energy, Redial, Dordrecht, 1121, (1986)
7. K. Wolf, H. Stanzl, A. Naumov, H.P. Wagnor, W. Kuhn, B. Kahn, W. Gebhardt, *J. Cryst. Growth* **138**, 412 (1994).
8. R.L. Gunshor, L.A. Kolodziejski, N. Otsuka, S. Datta, *Surf. Sci.* **174**, 522 (1986).
9. U. Pal, S. Saha, A. K. Chaudhari, V.V. Rao, H.D. Benarji, *J. Phys. D: Appl. Phys.* **22**, 965 (1989).
10. R. Sarma, N. Mazumdar and H.L. Das, *Bull. Mater. Sci.* **29**, 11 (2006).
11. G.I. Rusu, P. Prepilita, N. Apetroaei and G. Popa, *J. Optoelectron. Adv. M.* **7(2)**, 829 (2005).
12. H. Belakhder, A. Outzourhit, E.L. Ameziane, *Thin Solid Films* **382**, 30 (2001).
13. M. Neumann-Spallart, C. Konigstein, *Thin Solid Films* **265**, 33 (1995).
14. A.B. Kashyout, A.S. Arico, P.L. Antonucci, F.A. Mohamed, V. Antonucci, *Mater. Chem. Phys.* **51**, 130 (1997).
15. K.K. Mishra, K. Rajeshwar, *J. Electroanal. Chem.* **273**, 169 (1989).
16. M.C. Lin, P.Y. Chen, I.W. Sun, *J. Electrochem. Society* **149**, 653 (2001).
17. B. Bozzini, M.A. Baker, P.L. Cavallotti, E. Cerri, C. Lenardi, *Thin Solid Films* **388**, 361 (2000).
18. S. Venkatachalam, D. Mangalaraj, Sa.K. Narayandass, K. Kim, J. Yi, *Physica B* **358**, 27 (2005).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.47-52

PACS numbers: 81.20.Fw, 81.07.Bc, 81.40.Tv

LOW TEMPERATURE SYNTHESIS AND CHARACTERIZATION OF $ZnTiO_3$ BY SOL-GEL METHOD

***H. Beigi*[1], *V.H. Bindu*[2], *H.Z.R. Hamoon*[1], *K.V. Rao*[1]**

[1] Center for Nano Science & Technology, Institute of Science and Technology, JNTUH, Kukatpally, 500085, Hyderabad, India
E-mail: beigi.hasan@gmail.com

[2] Center for Environment, Institute of Science and Technology, JNTUH, Kukatpally, 500085, Hyderabad, India

$ZnTiO_3$ nanoparticles were prepared by a modified sol-gel method at the low sintering temperature of 550 ¤C. Titanium tetra isopropoxide and Zinc acetate dihydrate ($C_4H_{10}O_6Zn(H_2O)_2$) materials were used as a source of titanium and zinc, respectively. The prepared nanopowders were characterized by means of X-ray diffraction (XRD), thermal gravimetric analysis (TGA), field emission Scanning electron microscope (FE-SEM) and Raman spectroscopy. The XRD patterns and Raman spectra revealed that in the temperature range of 550 to 800 °C, $ZnTiO_3$ is the only zinc titanate compound exists in the samples.

Keywords: *ZINC TITANATE, $ZnTiO_3$, SOL-GEL METHOD, SEMICONDUCTOR, DIELECTRIC PROPERTIES.*

(Received 04 February 2011, in final form 20 March 2011)

1. INTRODUCTION

Dulin and Rase [1] and Bartram and Slepetys [2] reported that there are three compounds exist in the $ZnO-TiO_2$ system, namely, $Zn_2Ti_3O_8$ (cubic), $ZnTiO_3$ (hexagonal) and Zn_2TiO_4 (cubic). $Zn_2Ti_3O_8$ is a low temperature form of $ZnTiO_3$ existing below 820 °C [3]. At the temperature about 945 °C, $ZnTiO_3$ was reported to decompose into Zn_2TiO_4 (zinc orthotitanate) and TiO_2 [1]. Zn_2TiO_4 is stable up to its liquid temperature (1418 °C).

Zinc titanate ($ZnTiO_3$), a perovskite type oxide structure, has a promising material as a gas sensor [4] (for ethanol, NO, CO, etc.), paint pigment [5], catalyst [6] and etc. Also $ZnTiO_3$ has been reported as a material with an excellent electrical properties which could be a useful candidate as microwave resonator [7]. This material has a dielectric constant of 19, quality values of 30.0 GHz and the temperature coefficients of the resonant frequency of - 55 ppm/°C [8].

The structures of titanium dioxide (TiO_2), Zn_2 Ti_3O_8, $ZnTiO_3$ and Zn_2 TiO_4 comprised of TiO_6 octahedra. In rutile and in $ZnTiO_3$, the connection of the TiO_6 octahedra results chains and/or layers. As a result of this similarity, $ZnTiO_3$ is formed only in the presence of rutile [9].

Different methods have been reported to prepare $ZnTiO_3$ powder in the literatures, including conventional solid-state reaction [1], sol gel route [10], molten salt method [11]. The solid state reaction method has some

to Contents, page 1

disadvantages such as high firing temperature and difficulties to control the size of particles.

In this paper we prepared $ZnTiO_3$ nano crystalline powder at relatively low temperature by a modified sol-gel method.

2. EXPERIMENTAL PROCEDURE

2.1 Materials

All reagents were of analytical grade and used as received. Titanium tetra isopropoxide (TTIP) with a normal purity of 97 % and Ethanol (C_2H_5OH) were purchased from Sigma Aldrich. Zinc acetate dihydrate ($C_4H_{10}O_6Zn(H_2O)_2$) and Ethanolamine (C_2H_7NO) were obtained from Merck Chemicals, and 18 MΩ deionized water (H_2O) was used to prepare the solutions.

2.2 Preparation of Zinc Titanate powder

The Zinc titanate nano crystalline powders were prepared by a modified sol-gel method. The experimental procedures are as follows: 3.28 ml Ethanolamine (ETA) were dissolved in ethanol (50 ml). The solution was stirred for 10 min at room temperature following by addition of 7.8 ml TTIP. The solution was successively stirred at room temperature for 15 min (solution A). Meanwhile, 5.62 gr of Zinc acetate was dissolved in ethanol (20 ml) and sonicated for 15 min to prepare solution B. Solution B was subsequently added to solution A under vigorous stirring. The solution was subsequently stirred for further 30 min (solution C). 3.41 ml ETA and 1.86 ml deionized water were dissolved in ethanol (45 ml) under vigorous stirring for 15 min to prepare solution D. Subsequently, solution D added dropwise to solution C under stirring. The obtained sol was stirred for further 2 h and aged for 48 h at room temperature. As-prepared zinc titanate gel were dried at 80 °C for 24 h. The obtained solids were ground ant finally calcined at 550 and 800 °C for 2 h. (heating rate = 5 °C/min)

3. CHARACTERIZATION

The X-ray diffraction (XRD) patterns were obtained using an Inel diffractometer (XRG 3000, France), $\lambda = 1.78897$ Å. Structural characteristics of as-prepared powders were observed using field emission transmission electron microscopy (FE-SEM, CARL − ZEISS − Ultra tm 55). The structural evolution of samples was characterized by Raman spectroscopy (LabRAM HR, using a wavelength of 633 nm laser).

4. RESULTS AND DISCUSSION

4.1 Thermal analysis

TGA/DTA analysis was used to understanding the synthesis process. Fig. 1 shows TGA/DTA curves of dried gel powder heated in air at 50 °C/min α-alumina was used as the reference sample. By the fig. 1, the weight loss of dried gel takes place at four distinctly separable levels. During the first level, a small endothermic peak appeared at about 80 °C as shown in DTA curve, which indicated the loss of water and ethanol in the composite sol (9.5 % weight loss). Decomposition of the organic components and

to Contents, page 1

Ethanolamine takes place in the second level. The total weight loss in this level is about 26.6 %. The exothermic peak at 310 °C seems to be corresponded to the detachments of surface-modifier [12]. The third level following by 7 % weight loss in the range of 415 - 495 °C results from the dehydroxylation of Ti-OH into rutile − TiO_2, as corroborated by an exothermic peak at 425 °C. A dramatically peak at 540 °C with the weight loss about 13 % between 495 and 597 °C ascribed to the direct crystallization of $ZnTiO_3$ from an amorphous component. In the temperature range of about 600 to 800 °C there is no significant change in the weight of sample. The phase stability of ($ZnTiO_3$ composition) in the range of 600 to 800 °C confirmed by XRD patterns.

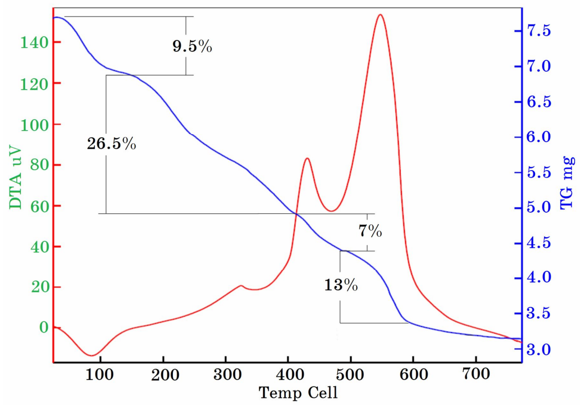


***Fig. 1** − The TGA/DTA curves of $ZnTiO_3$ gel*

4.2 XRD analysis

Powder X-ray diffraction (XRD) patterns of powders were recorded with an Inel diffractometer (XRG 3000, France), $\lambda = 1.78897$ Å using CoK_α radiation. As shown in fig. 2, the 2θ peaks appearing at 27.8, 38.2, 41.2, 47.3, 57.5, 62.9, 73.2 and 75.2 in the synthesized samples are attributed to the reflections from (110), (121), (110), (120), (220), (231), (130) and (211) plans of $ZnTiO_3$, respectively (JCPDS no. 850547). The rutile phase was detected by peak at $2\theta = 32.1$ (110), 42.3 (101), 48.5 (110), 64.4 (211) and 67.3 (67.3) (JCPDS no. 881172).

The crystallite size of $ZnTiO_3$ is determined by means of the Debye-Scherrer equation [13] (eq. 1) expressed as follows:

$$d = \frac{0,9\lambda}{\beta \cos\theta} \tag{1}$$

where d is the crystallite size, λ is the X-ray wavelength (1.78897 Å), β is the full width at the half maximum of the diffraction peak and θ is the Bragg diffraction angle. The crystallite sizes are about 41 and 61.8 nm for calcined samples at 550 and 800 °C, respectively.

to Contents, page 1

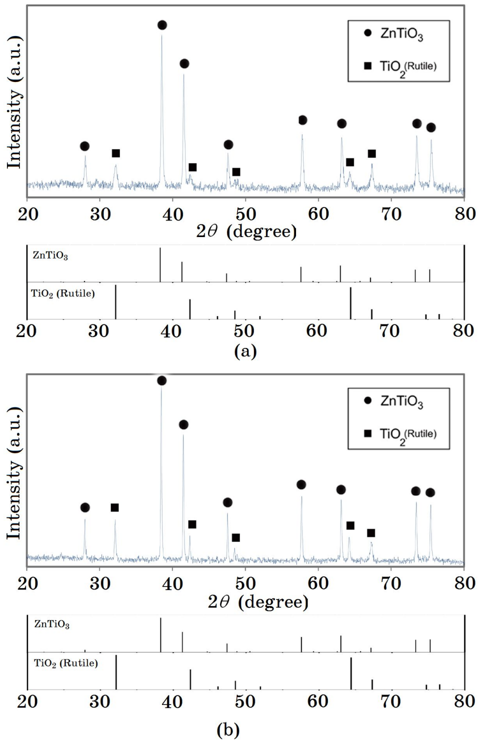


Fig. 2 *– XRD patterns of powders calcined at* 550 °C *(a) and* 800 °C *(b)*

4.3 SEM micrographs of $ZnTiO_3$ powders

Morphologies of the prepared powders were characterized by FE-SEM. Fig. 3a and 4a show the low magnification surface of prepared $ZnTiO_3$ and fig. 3b and 4b give the SEM micrographs of $ZnTiO_3$ powders after heat treatment at (3b) 550 and (4b) 800 °C for 2 h. Fig. 3b and 4b clearly show that a higher calcination temperature is an increase in grain size. These results are in good agreement with the grain sizes calculated by Debye-Scherrer equation.

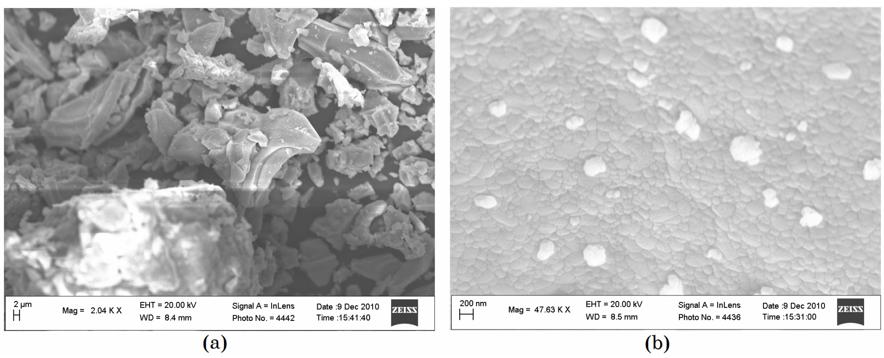


Fig. 3 *– Surface morphology of the powder calcined at* 550 °C

to Contents, page 1

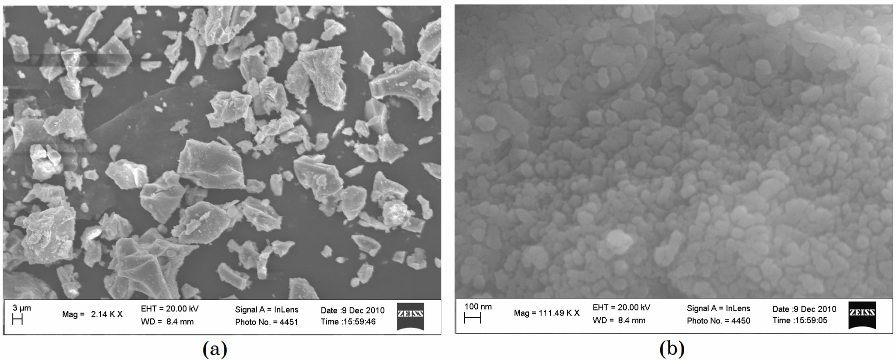


Fig. 4 *– Surface morphology of the powder calcined at* 800 °*C*

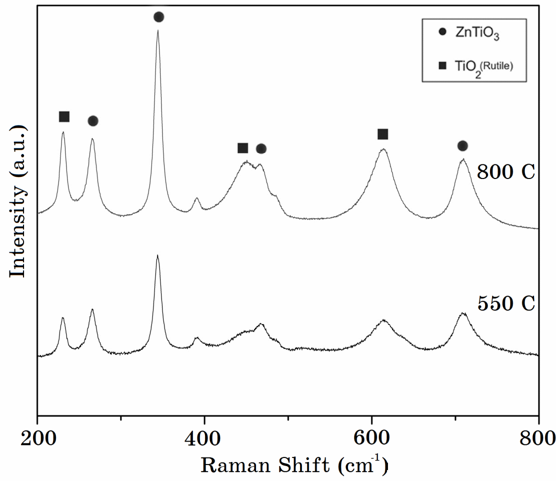


Fig. 5 *– Raman spectra of powders calcined at* 550 °*C and* 800 °*C*

4.4 Raman spectroscopy

To further investigate the synthesized powders, the Raman spectra were collected with 633 nm laser as excitation. Fig. 5 shows Raman spectra of calcined powders at 550 and 800 °C. The Raman spectroscopy results are in good agreement with XRD patterns. The peaks are located at about 264, 343, 462 and 705 cm^{-1} are ascribe to hexagonal $ZnTiO_3$. The peaks are located at about 229, 445 and 610 cm^{-1} related to existence of rutile in the samples [10].

5. CONCLUSION

$ZnTiO_3$ composite materials with nanocrystalline structure have been obtained using sol-gel method at relatively low temperature of 550 °C. Titanium tetra isopropoxide were used as titanium and zinc precursors. The obtained gel calcined under air at 550 and 800 °C. The grain sizes of $ZnTiO_3$ increase when calcination temperature increased.

to Contents, page 1

The authors would like to express their gratitude to Dr. S. Srinath and school of physics, university of Hyderabad for enabling use of SEM, XRD and Raman spectroscopy.

REFERENCES

1. F.H. Dulin, D.E. Rase, *J. Am. Ceram. Soc.* **43**, 125 (1960).
2. S.F. Bartram, R.A. Slepetys, *J. Am. Ceram. Soc.* **44**, 493 (1961).
3. O. Yamaguchi, M. Morimi, H. Kawabata and and K. Shimizu, *J. Am. Ceram. Soc.* **70**, 97 (1987).
4. H. Obayashi, Y. Sakurai, T. Gejo, *J. Solid State Chem.* **17**, 299 (1976).
5. A.T. McCord, H.F. Saunder, *U.S. Patent 2739019, Ceram. Abstr.* **24**, 155 (1945).
6. S.F. Bartram, R.A. Slepetys, *J. Am. Ceram. Soc.* **44**, 493 (1961).
7. H.T. Kim, S. Nahm, J.D. Byum, *J. Am. Ceram. Soc.* **82**, 3476 (1999).
8. H.T. Kim, S.H. Kim, S. Nahm, J.D. Byun, Y.H. Kim, *J. Am. Ceram. Soc.* **82**, 3043 (1999).
9. Z. Liu, D. Zhou, S. Gong, H. Li, *J. Alloy. Compd.* **475**, 840 (2009).
10. L. Hou, Y. Hou, M.K Zhu, J. Tang, J.B. Liu, H. Wang, H. Yan, *Mater. Lett.* **59**, 197 (2005).
11. X.R. Xing, C.Y. Zhang, L.J. Qiao, G.R. Liu, *J. Am. Ceram. Soc.* **89**, 1150 (2006).
12. X. Liu, M. Zho, F. Gao, L. Zhao, C. Tian, *J. Alloy. Compd.***450**, 440 (2008).
13. B.D. Cullity, S.R. Stock, *Elements of X-ray Diffraction* (London: Addison-Wesley: 1978).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.53-58

PACS numbers: 81.20.Fw, 81.07.Bc, 07.07.Df.

A SIMPLE SOL GEL PROTOCOL TOWARDS SYNTHESIS OF SEMICONDUCTING OXIDE NANOMATERIAL

***G.S. Devi*[1], *H.Z.R. Hamoon*[1], *A. Nanaji*[1], *K.R. Reddy*[1], *B. Sreedhar*[1], *J.V. Ramana Rao*[2]**

[1] Inorganic & Physical Chemistry division,
Indian Institute of Chemical Technology,
Tarnaka, Hyderabad, AP, India
E-mail: sarala@iict.res.in

[2] Center for Nano Science & Technology, Institute of Science and Technology, JNTUH,
Kukatpally, Hyderabad, 500085, AP, India

Nanostructured Tin oxide (SnO_2), powders was synthesized by employing a novel Sol-gel protocol at RT. A wide variety of techniques such as energy – dispersive spectroscopy(EDX), N_2 sorption, X-ray diffraction (XRD), have been used to study the formation process and characterization of the nanoparticles obtained. Transmission electron microscopy (TEM) has been applied to find out about the shape and size distribution of the particles. The nanoparticles thus synthesized were monodispersed, with an average particle size of ~ 10 nm and spherical in shape. The EDX analysis revealed the presence of Sn, O signal in the synthesized nanoparticles confirming the purity of the synthesized samples. This protocol appears promising for application in large-scale synthesis of nanoparticles.

Keywords: *SnO_2, NANOPARTICLES, SEMICONDUCTORS, XRD, TEM, BET.*

(Received 04 February 2011, in final form 17 March 2011)

1. INTRODUCTION

The size and shape dependent properties of nanomaterials provide a challenge to synthetic chemists for obtaining highly functional advanced materials. It is well known that shape and size of inorganic nanocrystals control their widely varying electrical, optical and catalytic properties [1-3]. Consequently, one of the emerging challenges in materials synthesis is achieving control over the morphology of nanocrystals. Since the discovery of mesoporous silica molecular sieves in 1992 [4-5] several supramolecular assembly pathways have been reported and extended to the synthesis of a variety of mesoporous metal oxide compositions [6]. Mesoporous materials have attracted considerable attention because of their remarkably large surface area and narrow pore size distribution, which make them ideal candidate for catalysts, molecular sieves, chemical sensors and as electrodes in solid state ionic devices. A number of related synthetic strategies, thermal evaporation [7], laser ablation [8], solution phase growth [9] and supra molecular templating methods [10] have been developed and a variety of materials, in terms of both composition and structure, have been prepared [11-13]. Among various metal oxides tin oxide is a wide band gap (E_g = 3.6 eV) semiconductor and has potential technological

to Contents, page 1

applicability in solid state gas sensors, [14] transparent conducting electrodes, [15] transistors solar cells and optical electronic devices [16]. The success in many of these applications relies critically on the preparation of crystalline SnO_2 with nanosize pore structure. Thus the objective of this study was to synthesize mesoporous SnO_2 stable at high temperature.

2. EXPERIMENTAL WORK

Nanosized tin oxide powder was synthesized using $SnCl_4.5H_2O$ as tin source and sodium dodecyl sulfate as surfactant. In a typical procedure 5 mmol of the surfactant was dissolved in deionized (DI) water (ρ = 18 MΩ) and then 1 M aqueous $SnCl_4$ solution was added to the above surfactant solution with stirring for 10 min, which resulted in a milky white suspension. After another 10 min. of stirring, the mixture was aged for 24 h at room temperature. The resulting product was centrifuged, washed with water and dried at 80 °C. The resultant tin oxide powders were calcined in the temperature range 400-700 °C for 2h at an heating rate of 5 °C min^{-1} crystal phase and crystallite size of SnO_2 were characterized by X-ray diffraction (XRD) (Siemens / 0-5000, $\Theta = 2\theta$) using CuKα radiation. Specific Surface area and pore size distribution were measured by a BET method using N_2 sorption isotherm determined at 77 K on a micromeritics pulse ASAP 2010 instrument The samples were degassed in a vacuum at 300 °C for 3h prior to measurement. Composition was analyzed by EDX analysis. Morphology of the SnO_2 powders was observed by transmission electron microscope (TEM; - Tech nai – 12 operated at 120 KeV).

3. RESULTS AND DISCUSSION

3.1 Characteristics of nanostructured SnO_2 powder EDS spectra

The composition of SnO_2 powder calcined at 400 °C was analyzed by EDS measurement as shown in Fig. 1.

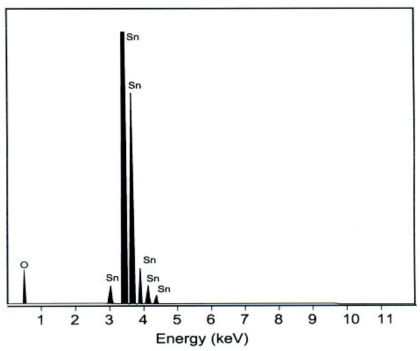


***Fig. 1** – EDS analysis of SnO_2 calcined at 400 °C*

The energy dispersive X-ray spectroscopy (EDS) analysis of nanoparticles dispersion confirmed the presence of Tin and oxygen signal no peaks of other impurity were detected.

to Contents, page 1

3.2 X-ray diffraction (XRD) spectra

The X-ray powder diffraction pattern of the as-synthesized nanoparticles and calcined at elevated temperatures are shown in Fig. 2

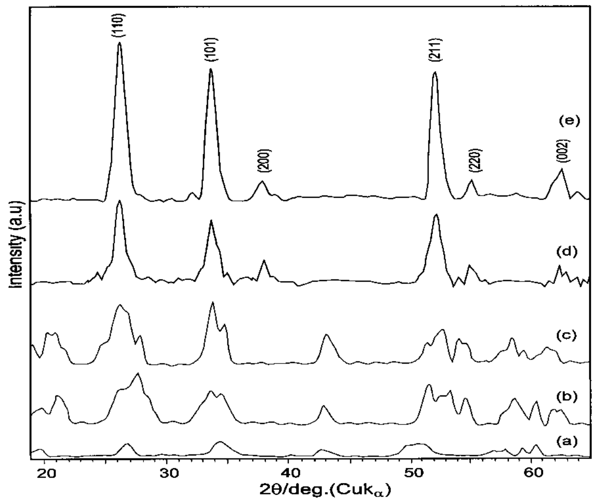


Fig. 2 *– XRD patterns of SnO_2 (a) as-synthesized, calcined at (b)* 400 °C *(c)* 500 °C *(d)* 600 °C *(e)* 700 °C

All peaks correspond to SnO_2 tetragonal structure and well match with the reported JCPDS data. The diffraction peaks became sharp with rise in calcination temperature, indicating the grain growth of SnO_2 crystallites. The crystallite size of each SnO_2 powder was calculated to be 4 - 9 nm using Scherrer's equation. Such change in crystallite size are obvious in TEM photographs shown in Fig. 4 c, indicating grain growth of SnO_2 particles, it is seen that sample calcined at 600 °C had a smaller particle size of ~ 10 nm. In accordance with these changes, the surface area decreased from 125 m^2g^{-1} (for as-synthesized powder) to 36.3 m^2g^{-1}(for powder calcined at 600 °C) as summarized in Table 1.

Table 1 *– Structural Properties of SnO_2 at diff. Calcination temperatures (T_c)*

T_c (°C)	Specific surface area (m^2g^{-1})	Pore volume[a] Dv/log d	Pore Size[b] (nm)	Crystalite Size[c] (nm)
As synthesized	125	0.81	4.2	5.3
400	121	0.79	4.5	6.3
500	61.2	0.65	8.0	6.9
600	36.3	0.45	10.0	7.9
700	35.8	0.17	19.0	7.9

[a,b] *Obtained from nitrogen adsorption.*
[c] *Estimated from XRD peak broadening 110 plane in Fig. 2*

According to Scherrer's formula the average SnO_2 crystallite size of about 5 nm in the (110) direction was derived from full width at half maximum (110) peak for as synthesized powder. The crystallite size has increased from 5 to 7.9 nm at elevated temperature (700 °C).

to Contents, page 1

3.3 Pore size distribution

From the pore size distribution (Fig. 3) it is seen that the pore volume decreased and pore diameter at maximum pore volume increased with rise in

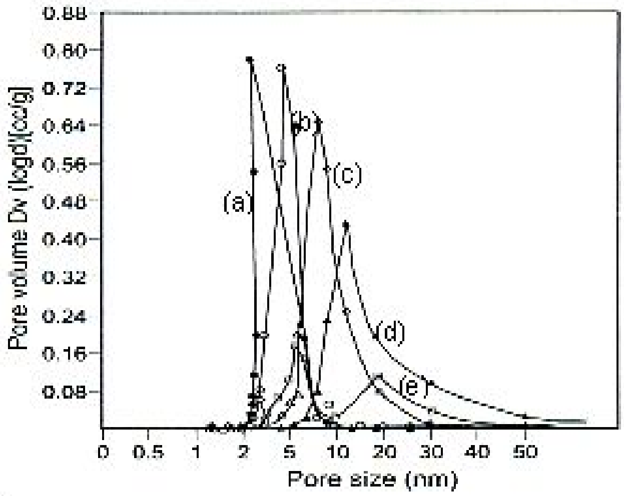


Fig. 3 *– The Barret-Joyner-Halenda pore size distribution determined N_2 adsorption (a) as- synthesized, calcined at (b)* 400 °C *(c)* 500 °C *(d)* 600 ° C *(e)* 700 °C

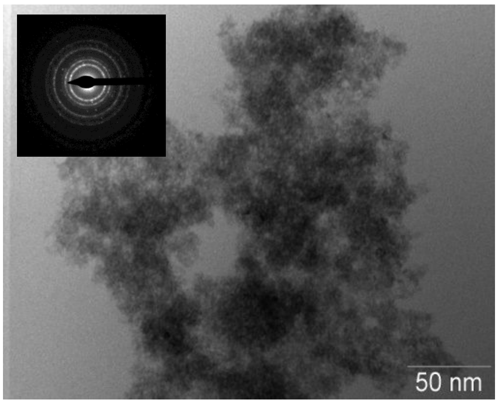
calcination temperature. The specific surface area (SSA) was significantly high when SnO_2 powder was fired at 400 °C, the change in specific surface area and crystallite size with calcination temperature of SnO_2 powder was also determined. High surface area may result from the breakup of large agglomerates into small crystallites, exploring more surfaces. It is seen that the specific surface area decreased and simultaneously particle size increased with increase in calcination temperature (T_c).

3.4 TEM micrograph

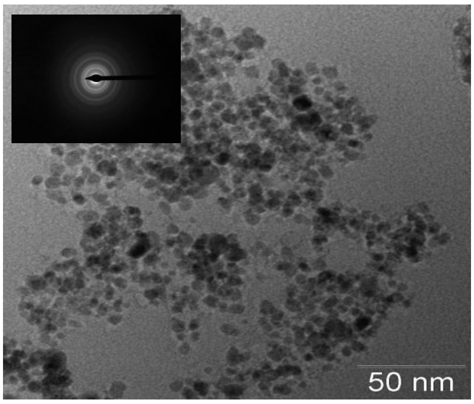
The surface morphology and particle size of SnO_2 nanoparticles were investigated Transmission Electron Microscopy studies.

TEM analysis indicated that the calcined sample were in the nanometer range as seen from Fig. 4. It is observed that, as synthesized samples are mesoporous with particle size of ~ 5 nm, (Fig. 4a). After firing at 400 °C for 2 h the particle size still remains ~ 7 nm (Fig. 4b) with increase in temperature to 600 °C the size increased to ~ 10 nm (Fig. 4c). It is worth noting that even after calcination at 600 °C the particle size still remained small. The inset of the figure shows the SAED pattern, as synthesized sample is mesoporous and calcined samples are clearly polycrystalline.

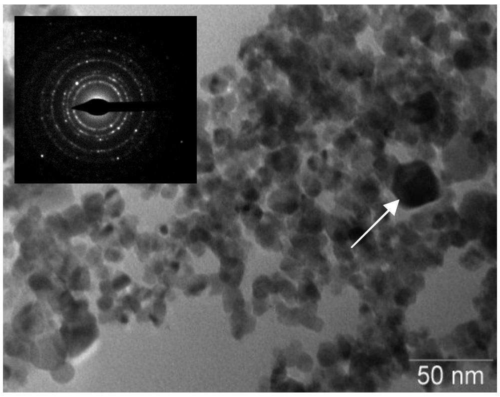
to Contents, page 1



a



b



c

***Fig. 4** – TEM micrograph of SnO_2: as synthesized* 400 °C *(a, b) and* 600 °C *(c). The inset shows the selected area electron diffraction (SAED) pattern*

4. CONCLUSION

In conclusion, simple surfactant templated process has been adopted for the synthesis of SnO_2 nanosized powder, having high surface area (121 m^2g^{-1} for powder calcined at 400 °C) and small grain size (~ 7 nm). And these nanoparticles were thermally stable with tetragonal structure. EDS analysis of SnO_2 powder showed the presence of Sn and O as only detected elements.

ACKNOWLEDGEMENT

Financial support by department of Science and Technology (DST, India) is gratefully acknowledged. A. Nanaji and Ravindranath Reddy thank CSIR and UGC for research fellowship.

REFERENCES

1. A.P. Aliviastos, *Science* **271**, 933 (1996).
2. T. Huang, R.W. Murray, *J. Phys. Chem. B* **105**, 12498 (2001).

to Contents, page 1

3. P.V. Kamat, *J. Phys. Chem. B* **106,** 7729 (2002).
4. C.T. Kresge, M.E. Leonowicz, W.J. Roth, J.C. Vartuli, J.S. Beck, *Nature* **359,** 710 (1992).
5. J.S. Beck, J.C. Vartuli, W.J. Roth, M.E. Leonowicz, C.T. Kresge, K.D. Schmitt, C.T.W. Chu, D.h. Olson, E.W. Sheppard, S.B. McCullen, J.B. Higgins, J.L. Schlenker, *J. Am. Chem. Soc* **114,**10834 (1992).
6. P. Behrens, *Angew. Chem. Int. Ed. Ehg.* **35,** 515 (1996).
7. Z.W. Pan, Z.R. Dai, Z.L. Wang, *Science* **291,** 1947 (2001).
8. (a) J. Hu, Y. Bando, Q. Liu, D. Golberg, *Adv. Func. Mater.* **13,**493 (2003). (b) Z. Liu, D. Zhang, S. Han, C. Li, T. Tang, W. Jin, X. Liu, B. Lei, C. Zhou, *Adv. Mater.* **15,** 1754 (2003).
9. D.F. Zhang, L.D. Sun, J.L. Yin, C.H. Yan, *Adv. Mater.* **15,** 1022 (2003).
10. C.R Martin, *Science* **266,** 1961 (1994).
11. Q. Huo, D.I. Margolese, U. Ciesla, D.g. Demuth, P. Feng. T.E. Goe, P. Sieger, A. Firouzi, B.F. Chmelka, F. Schuth and G.D. Stucky, *Chem. Mater.* **6,** 1176 (1994).
12. A. Sayari and P. Liu, *Microporous Mater.* **12,** 149 (1997).
13. J.Y. Ying, C.P. Mehnert and M.S. Wong, *Angew. Chem. Int. Ed*. **38,** 56, (1996).
14. (a) M. Law, H. Kind, B. Messer, F. Kim, P.D. Yang, *Angew. Chem. Int. Ed* **41,** 2405 (2002).
    (b) Y. Wang, X. Jiang, Y. Xia, *J. Am. Chem. Soc.* **125,** 16176 (2003).
15. Y.S. He, J.C. Campbell, R.C. Murphy, M.F. Arendt, J.S. Swinnea, *J. Mater. Res.* **8**, 3131 (1993).
16. (a) P.G. Harrison, M.J. Willet, *Nature* **332,** 337 (1988).
    (b) S. Ferrere, A. Zaban, B.A. Gregg, *J. Phys. Chem. B* **101,** 4490 (1997).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.59-66

PACS numbers: 81.20.Fw, 81.07.Bc, 07.07.Df.

SYNTHESIS AND CHARACTERIZATION OF Co-DOPED SnO_2/TiO_2 SEMICONDUCTOR NANO CRYSTALLITES VIA SOL-GEL METHOD

***H.Z.R. Hamoon*[1], *G.S. Devi*[1], *H. Beigi*[2], *J.V.R. Rao*[2], *K.R. Reddy*[1], *A. Nanaji*[1]**

[1] Inorganic and Physical Chemistry division,
Indian Institute of Chemical Technology,
Tarnaka, Hyderabad, AP, India
E-mail: hedayatzadehamoon@yahoo.com

[2] Center for Nano Science & Technology,
Institute of Science and Technology, JNTUH,
Kukatpally, Hyderabad, 500085, AP, India

SnO_2/TiO_2 nano particles are novel wide band gap semiconductors with modified applications of SnO_2 and TiO_2 in some fields including gas sensing, photo catalytic, solar cells and so on. The Co-doped SnO_2/TiO_2 nano particles were obtained via sol-gel method with different amounts of doping material as 2.5 %, 6 % and 10 mol %. The crystallite sizes of resulting material were from 3.8 nm for 0.1 wt % Co-doped SnO_2/TiO_2 to 19.1 nm for un-doped. Morphology and nanostructure of the crystalline SnO_2/TiO_2 nano particles were characterized by means of X-ray diffraction, Raman spectroscopy, Fourier transform infrared spectroscopy (FTIR), Thermal gravimetric analysis (TGA), field emission scanning electron microscopy (FESEM) and energy dispersive X-ray spectroscopy (EDX). It has been shown that fine semiconductor nano structures were formed.

Keywords: *SnO_2/TiO_2, Co-DOPED, SEMICONDUCTOR, XRD, NANOPARTICLE.*

(Received 04 February 2011, in final form 20 March 2011)

1. INTRODUCTION

SnO_2 is an opto electrical n-type semiconductor that has many interesting properties. These properties have made it a reliable material for gas sensing and catalysis applications [1, 2]. Even gas sensors based on SnO_2 materials are commercially available today [3]. SnO_2-based nano powders and devices are obtained by means of a variety of synthesis techniques including the co precipitation [4], ion sputtering [5], microwave heating method [6], surfactants mediate [7,8], sol-gel methods [9] and so on.

On the other hand TiO_2 is also an n-type semiconductor that has been subject of many research studies in the last decade. It has been extensively employed in many applications mainly to photo catalytic decontamination treatments. [11-12].

One of the effective methods for obtaining desired material with modified properties is mixing two different semiconductors with appropriate conduction and valence band edges [13]. With coupling two metal oxide nanoparticles, better structures with more advantages for sensor applications can be achieved [14-15].

The energy levels of the valence and conduction bands are 3.7 eV and (ECB = 0 V versus NHE at pH 7) for SnO_2 and 2.7 eV (ECB = – 0.5 V versus

to Contents, page 1

NHE at pH 7) for TiO_2 [16, 17]. The band gap energy level of SnO_2 is higher than that of TiO_2; the conduction band of SnO_2 is at a lower level than that TiO_2's. The higher reduction power of electrons and the higher oxidation power of holes correspond to the higher position of conduction band and the lower position of valence band, respectively. Therefore, mixing two semiconductors with different energy levels for their corresponding conduction and valence bands can provide an approach to achieve better applications by increasing the efficiency of charge separation, charge carrier lifetime, ,interfacial charge transfer rate and extending the energy range of photo excitation [16].

Mixed oxides can be formed in three ways along follow lines:

(1) Chemical compound.
(2) Solid solution
(3) Mix of (1) and (2) types.

The SnO_2-TiO_2 mixture belongs to second category [18]. The SnO_2-TiO_2 coupled semiconductor commonly is synthesized by mixing the colloidal solutions of SnO_2 and TiO_2 [19, 20].In recent years the properties of Co-doped SnO_2 nanoparticles and Co-doped TiO_2 nano particles have been exhaustively investigated. [21, 22, 23]

In the case of Co-doped TiO_2 the main focus has been on the inherent ferromagnetism property of resulting nano structure at or above room temperature. Co: TiO_2 nano particles are one of the most prominent diluted magnetic semiconductors (DMS) In the field of spintronics, especially for magneto-opto-electronic applications [24-25] DMSs have significant potential for future applications in the solar cells, biomedical and environmental fields. Co-doping SnO_2 also has been widely studied [21]. One of the new emerging applications of SnO_2 nano particles doped with Co is in varistor base materials.

SnO_2 with native oxygen vacancies compensated by electrons:

$$O_0^x = 2e + V_0'' + 0.5O_{2(g)}$$

Using tin oxide ceramics is limited as dense ceramic since it is hard to make it denser because of evaporation and condensation, which enhance the grain growth, mostly dominate mass transport [26, 27]. Several sintering aids have been used to improve the densification of SnO_2 ceramics [23, 28]. Recent reports suggest Co is used as sintering aids [23, 29].

Hence the investigation of nano particles of mix of SnO_2-TiO_2 doped via Co would have many advantages and it selected as a novel semiconductor for the current subject of this study. In this paper, a fine mixture nano-structural SnO_2-TiO_2 with different dosages of doped-Co is synthesized via sol-gel method while $SnCl_4$ as a common precursor for producing SnO_2 nano crystallite structure has been used. The mixture ratio of 3:1 has been kept constant for all samples and the Co doped amount varied from 0 % to 10 mol %. Geometry, particle size, morphology of nano crystallite structures are estimated by measurements.

2. EXPERIMENTAL WORK

Stannic chloride ($SnCl_4$, spectrochem Pvt. Ltd, Mumbai), titanium chloride ($TiCl_4$, spectrochem Pvt. Ltd, Mumbai), Poly Ethylene Glycol (6000LR, s-d fine-chem Limited, Mumbai), Ammonia solution 25% (Extra pure, s-d fine-

to Contents, page 1

chem Limited, Mumbai) were purchased as precursors. 0.3 mol ($SnCl_4$) and 0.1 mol $TiCl_4$ were separately diluted using distilled water under simultaneous stirring condition for 45 min. 13 grams of PEG was dissolved in 300 ml distilled water and was stirred for 30 min until transparent solution was obtained. Diluted $SnCl_4$ and $TiCl_4$ were added drop by drop to the above solution simultaneously. Solution continued to remain under vigorous stirring for more than an hour. Appropriate amounts of $CoCl_2$ for (0.25 mol %) Co-doping after dissolving in distilled water was added to mixture and was stirred for 30 min. 25 % aqueous ammonia solution was employed for adjusting PH value of the resulting solution to around 3. Then stirring condition was continued vigorously for one hour. The resulting sol was filtered and washed with distilled water and acetone twice, respectively, and dried overnight at 80 °C. Resulting solid was ground followed by calcination at 550 °C for 2 hours. This procedure was repeated three times with varying the amounts of $CoCl_2$ addition for obtaining four different nano powders with 0%, 2.5%, 6%, 10 mol % Co-doped finally.

3. RESULTS

3.1 TGA

Main Co-doped SnO_2/TiO_2 (6 mol % doped) thermo gravimetric analysis (TGA) was accomplished in a flow of air with a temperature ramp of 50 °C/min and α-alumina was used as the reference. Fig. 1 shows the weight loss (TG) and the differential thermo-analysis (DTA) curves corresponding to Co-doped SnO_2/TiO_2 without any thermal annealing. The TG curve exhibits one endothermic weight loss at temperatures lower than 130 °C with 16.6 % weight loss that is associated with the loss of the residual water and trapped solvent in the particles. From 130 °C to 275 °C and 275 °C to 310 °C there are two small exothermic peaks corresponding to 6.6 % and 9.3 % relative weight losses respectively. About 360 °C a vigorous exothermic peak was observed accompanied by 9.3 % loss weight. Last exothermic peak at 500 °C corresponds to 6.6 weight loss. These exothermic peaks are likely due to decomposition of resulting nano crystallite structures, as confirmed by XRD analysis shown in Fig. 2. No more peak and no further weight loss observed in the TGA/DTA curves.

3.2 XRD

Structural properties of all 0 %, 2.5 %, 6 %, 10 mol % Co-doped SnO_2/TiO_2 samples were studied using X-ray diffraction (XRD: XRG 3000, Inel, France). The XRD results obtained were compared to the Joint Committee on Powder Diffraction Standards (JCPDS) X-ray data file. The crystallite size of the samples was calculated using the Scherrer formula:

$$T = 0.9\lambda/(\beta\cos\theta)$$

Where T is the particle size in nanometers, λ is the wavelength of Co radiation ($\lambda = 1.78897$), β is the FWHM of the strongest peak and θ is the peak position. By using the equation the crystal size of all different nanoparticles were calculated for each calcined temperature and are depicted

to Contents, page 1

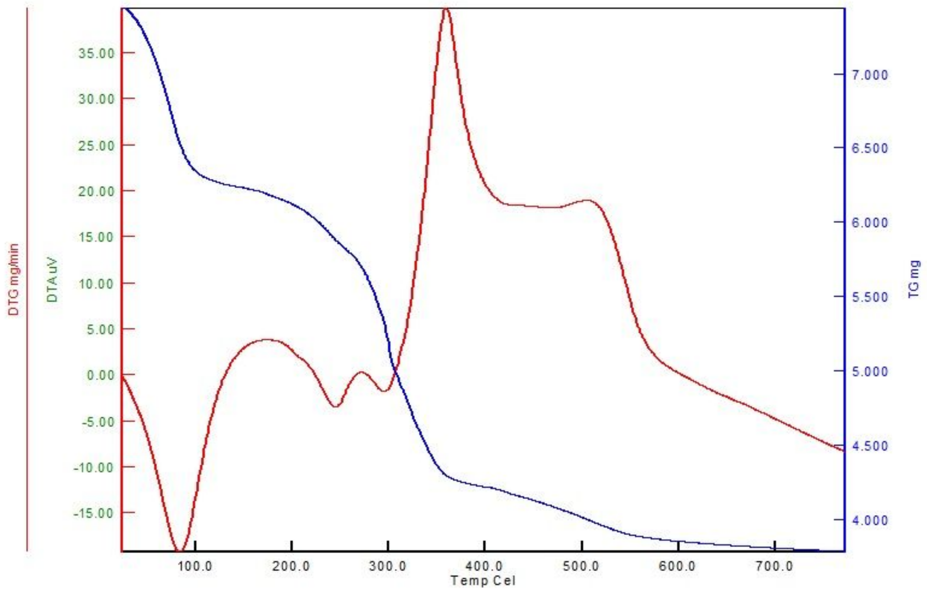


Fig. 1 *– TGA and DTA curves of Co-doped SnO_2/TiO_2*

in Table 1. Presence of SnO_2 and rutile TiO_2 were confirmed by PDF# 77-0451 (5 peaks are match) and PDF#76-0326 (2 peaks 111 and 211) respectively. Also CoO (hexagonal) peak was observed (PDF#89-2803 and peak 110). Decreasing the crystallite size is proportional with amount of Co-doped to SnO_2/TiO_2.

Table 1 *– Crystallite and particle size of Co-doped SnO_2/TiO_2 nanostructures*

Sample	Crystallite size (nm)	Particle size (nm)
Un-doped SnO_2/TiO_2	19.1	27
2.5 mol % Co-doped SnO_2/TiO_2	15.5	21
6 mol % Co-doped SnO_2/TiO_2	4.3	17
10 mol % Co-doped SnO_2/TiO_2	3.8	16

3.3 Raman

The Raman spectra of all samples were depicted in Fig. 3 and were compared with literature data [30, 31, 32]. For characterization the wave length 633 cm^{-1} was used for adjusting Raman apparatus (Labram Horibb yvon). Commercial CoO Raman spectra are shown in Fig. 3e, two peaks at 648, 485 cm^{-1}. The SnO_2 peak about 630 cm^{-1} corresponding to A_{1g} mode, and two bands at 773 and 472 cm^{-1} corresponding to B_{2g} and E_g modes, respectively; this being in accordance with literature data. TiO_2 rutile has peaks with following values: 235, 447 [30] or 235, 447, 612 [31].

Peaks of TiO_2 rutile and SnO_2 and CoO were confirmed and that is in agreement with other characterization data.

to Contents, page 1

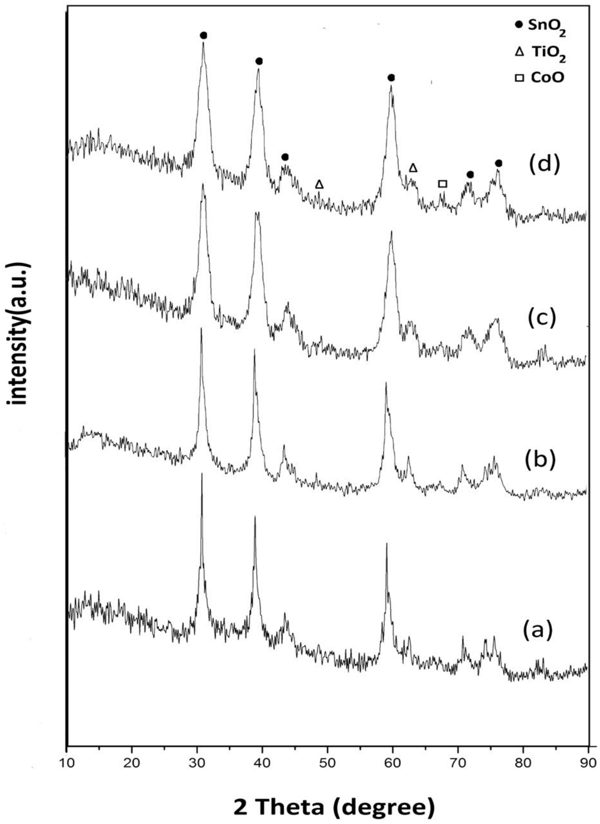


Fig. 2 *– XRD pattern of Co-doped SnO_2/TiO_2 (a) un-doped (b)* 2.5 *mol % doped (c)* 6 *mol % doped (d)* 10 *mol % doped*

3.4 FTIR

The FTIR analysis was carried out in order to determine the functional group of materials existed in Co-doped to SnO_2/TiO_2 samples. The analysis was recorded using FTIR spectrometer (bruker) in a range between 4000 and 400 cm^{-1}.

Fig. 4 shows the FTIR spectra of Co-doped to SnO_2/TiO_2. The bands at 3543 - 3393 cm^{-1} correspond to the O-H mode of vibration [33]. The broad O-H peaks become narrower with an increase Co-doped amount, 3380 and 3050 cm^{-1}, which may be due to the adsorbed water and NH. The N-H band observed around ca. 3050 cm^{-1} may be due to the use of ammonia to promote $SnCl_4$ and $TiCl_4$ hydrolysis, which leads to the formation of ammonia. The strong asymmetric stretching mode of vibration of C = O was observed between 1668 cm^{-1}. The symmetric stretching occurs between 1400 and 1334cm^{-1} because of the presence of C-O. The C-O-C peak is present about 1070 cm^{-1}. According to Du et al. [34], the C-O-C peak usually appears at 1256 cm^{-1}. The absorption peaks at 1359, 1343, 1280, 1240, 1149 and 1077 cm^{-1} result from the bending vibration of the O-H bond and stretching vibration of the C-O bond of the -CH_2-OH group. In fact the XRD and EDX analysis proved that the synthesized powders in this work are Co-doped to SnO_2/TiO_2 with the good stoichiometric composition.

to Contents, page 1

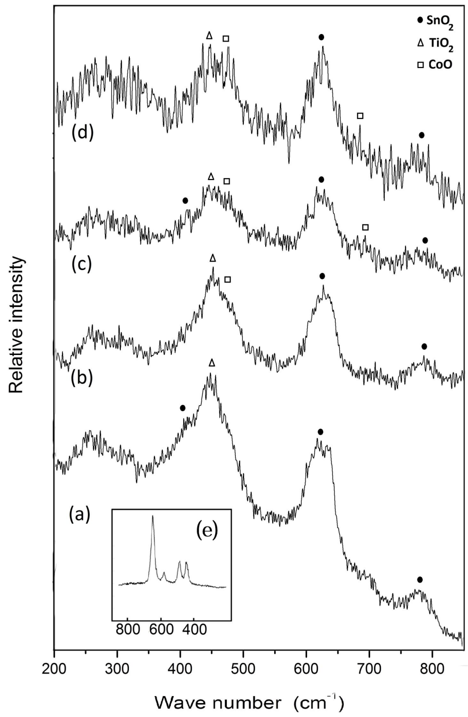


Fig. 3 – *Raman spectra of Co-doped SnO_2/TiO_2 un-doped (a), 2.5 mol % doped (b,) 6 mol % doped (c), 10 mol % doped (d), and commercial CoO (e)*

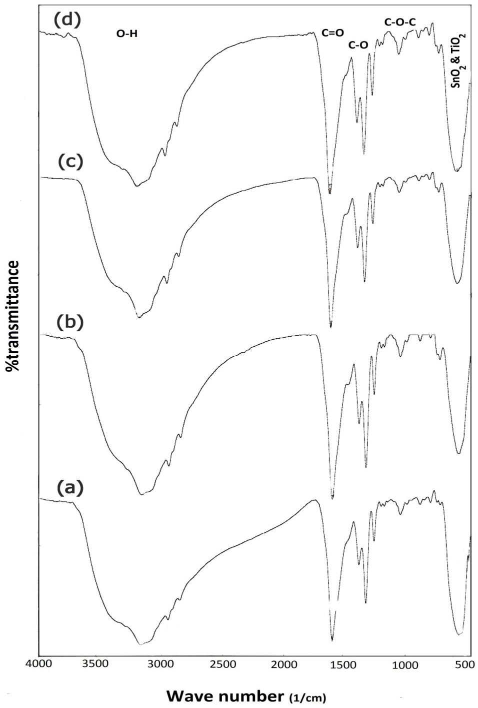


Fig. 4 – *FT-IR spectra of Co-doped SnO_2/TiO_2 un-doped (a), 2.5 mol % doped (b), 6 mol % doped (c), and 10 mol % doped (d)*

The particle size caused a large shift in the IR peak. The SnO_2 peak that appears 660 cm^{-1} shows that formation of tin oxide was completed. The FTIR results support the RAMAN, XRD, FESEM results.

3.5 FESEM

The surface morphology, particles size and composition of Co-doped to SnO_2/TiO_2 nanoparticles were investigated by field emission scanning electron microscopy (FESEM) and energy dispersive X-ray spectroscopy (EDX). FESEM and EDX images of Co-doped to SnO_2/TiO_2 powder nanoparticles are shown in Fig. 5. Nano particles with relatively high agglomeration are shown in Fig. 5 a and b. This agglomeration is a result of the acidic sols during the synthesis process. The particle sizes were estimate for all samples using images Table 1.

to Contents, page 1

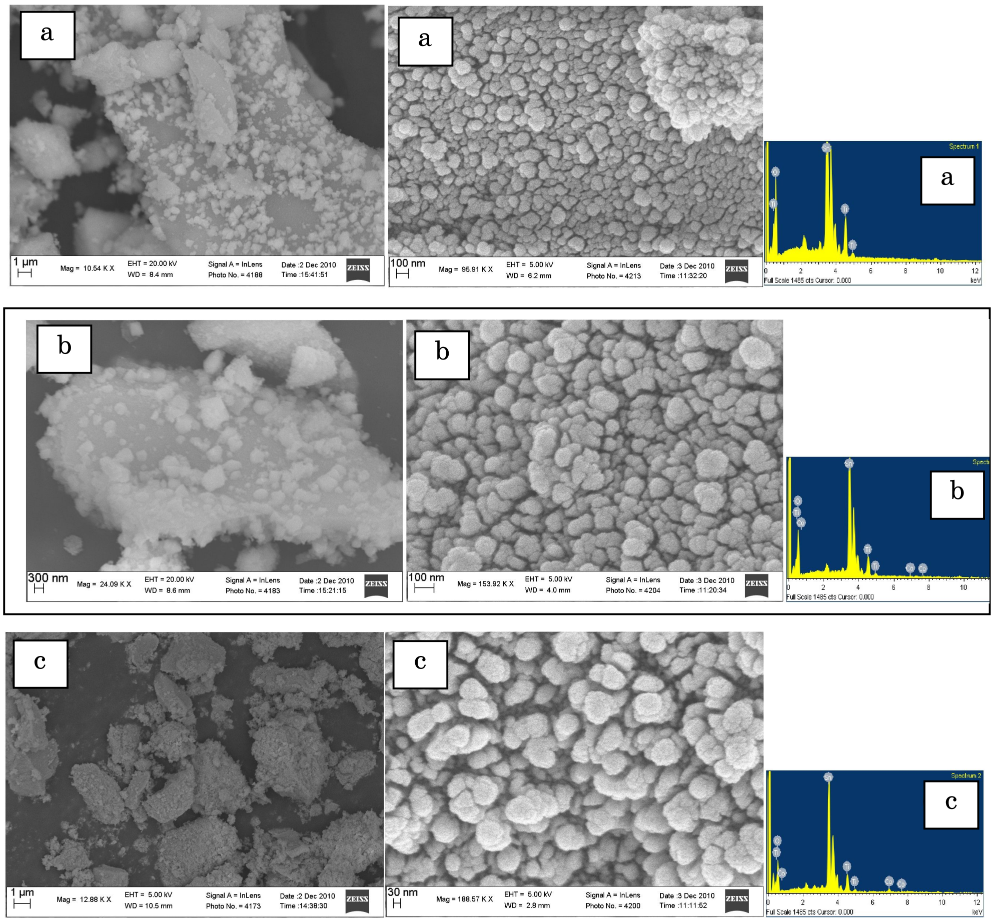


***Fig. 5** – FESEM and EDX images of Co-doped to SnO_2/TiO_2 un-doped (a) 2.5 mol % doped (b) 10 mol % doped (c)*

4. CONCLUSION

A modified sol-gel process was designed to prepare Co-doped to SnO_2/TiO_2. The products were characterized by means of X-ray diffraction, Raman spectroscopy, Fourier transform infrared spectroscopy (FTIR), Thermal gravimetric analysis (TGA), field emission scanning electron microscopy (FESEM) and energy dispersive X-ray spectroscopy (EDX). Crystallite and particle size of the nano structures were decreased from 19.1 to 3.8 nm and 27 to 16 respectively by the addition of Co to the precursor solution. Fine Co-doped to SnO_2/TiO_2 nano particles were obtained and applications were discussed.

ACKNOWLEDGEMENT

Hamoon H.Z.R. thanks Director IICT for permitting him to carry out his thesis work at IICT. A. Nanaji and Ravindranath Reddy thank CSIR and UGC for research fellowship.

to Contents, page 1

REFERENCES

1. M. Batzill, U. Diebold, *Prog. Surf. Sci.* **79,** 47 (2005).
2. M. Tiemann, *Chem. Eur. J.* **13,** 8376 (2007).
3. Microsens Inc. (http://www.microsens.ch).
4. A. Mosquera, J.E. Rodrнguez-Pбez, J.A. Varela, P.R. Bueno, *J. Eur. Ceram. Soc.* **27,** 3893 (2007).
5. M. Ruske, G. Brauer, J. Pistner, U. Pfafflin, J. Szczyrbowski, *Thin Solid Films* **351,** 146 (1999).
6. J.-J. Zhu, J.-M. Zhu, X.-H. Liao, J.-L. Fang, M.-G. Zhou, H.-Y. Chen, *Mater. Lett.* **53,** 12 (2002).
7. J. Pal, P. Chauhan, *Mater. Charact.* **60,** 1512 (2009).
8. C. Liang, Y. Shimizu, T. Sasaki, N. Koshizaki, *J. Phys. Chem. B* 107, 9220 (2003).
9. A. Kurz, K. Brakecha, J. Puetz, M.A. Aegerter, *Thin Solid Films.* **502,** 212 (2006).
10. N. Negishi, T. Iyoda, K. Hashimoto, A. Fujishima, *Chem. Lett.* **9,** 841 (1995).
11. I. Sopyan, M. Watanabe, S. Murasawa, *Chem. Lett.* **1,** 69 (1996).
12. A.L. Linsebigler, G.Q. Lu, J.T. Yates, *Jr. Chem. Rev.* **95,** 735 (1995).
13. J. Shang, W. Yao, *Appl. Catal. A-Gen* **257,** 25 (2004).
14. K. Zakrzewska, *Thin Solid Films* **391,** 229 (2001).
15. P. Nelli, L.E. Depero, M. Ferroni, S. Groppelli, V. Guidi, F. Ronconi, L. Sangaletti, G. Sberveglieri, *Sens. Actuat. B Chem.* **31,** 89 (1996).
16. O. Carp, C.L. Huisman, A. Reller, *Prog. Solid State Chem.* **32,** 33 (2004).
17. K. Vinodgopal, P.V. Kamat, *Environ. Sci. Technol.* **29,** 841 (1995).
18. N.V. Duy, N.V. Hieu, *Physica E* **41,** 258 (2008).
19. P.V. Kamat, *Chem. Rev.* **93,** 267 (1993).
20. I. Bedja, S. Hotchandani, P.V. Kamat, B. Bunsen-Ges, *Phys. Chem. Chem. Phys.* **101**, 1651 (1997).
21. R. Parra, L.A. Ramajo, *Mater. Res. Bull.* **43,** 3202 (2008).
22. K.A. Griffin, A.B. Pakhomov, *Phys. Rev. Lett.* **94,** 157204 (2005).
23. R. Metz, D. Koumeir, J. Morel, J. Panisot, M. Houabes, M. Hassanzadeh, *J. Eur. Ceram. Soc.* **28,** 829 (2008).
24. H. Ohno, *Science* **281,** 951 (1998).
25. S.J. Pearton, W.H. Heo, M. Ivill, D.P. Norton, T. Steiner, *Semicond. Sci. Technol.* **19,** 59 (2004).
26. C. Li, J. Wang, W. Su, H. Chen, W. Zhong, P. Zhang, *Ceram. Int.* **27,** 655 (2001).
27. T. Kimura, S. Inada, T. Yamaguchi, *J. Mater. Sci.* **24**, 220 (1989).
28. J.A Varela, O.J. Whittemore, E Longo, *Ceram. Int.* **16,** 177 (1990).
29. S.A. Pianaro, P.R. Bueno, E. Longo, J.A. Varela, *J. Mater. Sci. Lett.* **14,** 692 (1995).
30. V.N. Kuznetsov, N. Serpone, *J Phys. Chem. B.* **110,** 25203 (2006).
31. T. Hirata, K. Ishioka, M. Kitajima, H. Doi, *Phys. Rev. B.* **53,** 8442 (1996).
32. Y. Liu, M. Liu, *Adv. Funct. Mater.* **15,** 57 (2005).
33. D.M. Fernandes, R. Silva, A.A.W. Hechenleitner, E. Radovanovic, M.A.C. Melo, E.A.G. Pineda, *Mater. Chem. Phys.* **115,** 110 (2009).
34. H. Du, F. Yuan, S. Huang, J. Li, Y. Zhu, *Chem. Lett.* **33,** 770 (2004).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.67-84

PACS numbers: 77.84.Bw, 71.20.Nr, 61.66.Fn, 73.61.Ey, 81.05.Ea, 61.72.Lk, 61.72.Ff

A STRATEGIC REVIEW OF REDUCTION OF DISLOCATION DENSITY AT THE HETEROGENIOUS JUNCTION OF GAN EPILAYER ON FOREIGN SUBSTRATE

***S.Das Bhattacharyya*[1], *P. Mukhopadhyay*[1], *P. Das*[1], *D. Biswas*[2]**

[1] Indian Institute of Technology,
Advanced Technology Development Centre, IIT Kharagpur, 721302, India
E-mail: papri21@gmail.com

[2] Indian Institute of Technology
Dept. of E&ECE, IIT Kharagpur, 721302, India

Now-a-days for long range microwave communication, especially for space applications, devices capable to operate at a high power and high frequency are desired. Compound Semiconductor (CS), mainly Gallium Nitride (GaN) based heterostructure electronic devices are the only available solutions till now to fulfil these criteria. However, looking from a cost and manufacturing perspective, GaN substrate has considerable drawbacks like non-availability, expense as well as compulsion to use older technologies for device designing as the wafer diameter is small. A potential solution for performance/cost dilemma is to grow high quality GaN as active layer on a well matured substrate by metamorphic technique. Metamorphic buffer technology allows the device designer an additional degree of freedom to optimize the transistor at high frequency for high gain and power applications. But this metamorphic buffer technology has some drawbacks, too. The main limiting factor for this technology is the propensity to develop dislocation at the heterojunction due to lattice mismatch between the grown layer and the substrate. A good quality metamorphic buffer can only be achieved by reduction of dislocation density at the heterojunction. This paper reviews the progress being made towards reduction of dislocation density of GaN based devices grown on Silicon Carbide (SiC), Sapphire (Al_2O_3) and Si substrate, respectively, in terms of material parameters and growth issues.

Keywords: *GALLIUM NITRIDE, SILICON CARBIDE, SAPPHIRE, DISLOCATION DENSITY, EPITAXIAL GROWTH, CRYSTAL STRUCTURE.*

(Received 04 February 2011, in final form 18 March 2011)

1. INTRODUCTION

To improve carrier transport in GaN films as well as to improve the efficiency and to reduce the cost of GaN-made optoelectronic devices, reduction of threading dislocation density is the most important [1, 2]. The main reason of dislocation density is lattice mismatch between the substrate and the GaN layer. Moreover, material's crystal structure, surface finishes, composition, reactivity, chemical, thermal and electrical properties are also considered for choosing a substrate-material for GaN growth (since the substrate properties are ultimately responsible for the efficiency of GaN-made devices) [3]. Some of the above mentioned substrate properties are considered to reduce the dislocation density of GaN.

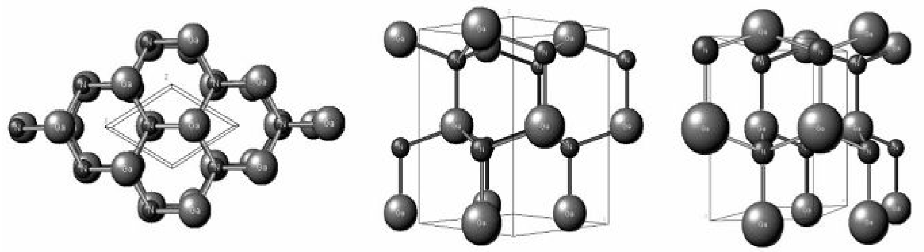
to Contents, page 1

Dislocation free GaN devices can be realized by growing the same on GaN substrate, as there will be no lattice mismatching. Defect-free bulk GaN can be grown at very high temperature (1400 - 1600 °C) and pressure (15 - 20 kbar) on GaN substrate, but it took very large growth time [4]. Moreover, bulk GaN substrate is not commercially available. The resultant of this fact is that the researchers are tending towards the growth of GaN film on foreign (e.g., Si, SiC Al_2O_3 (sapphire) etc.) substrates, called heteroepitaxial growth of GaN film. Heteroepitaxial growth is highly dependent on the properties of the substrates – both the inherent properties (lattice constant, thermal expansion co-efficient) and process-induced properties (e.g., surface roughness, step height, etc.) [3].

2. PROPERTIES OF GAN

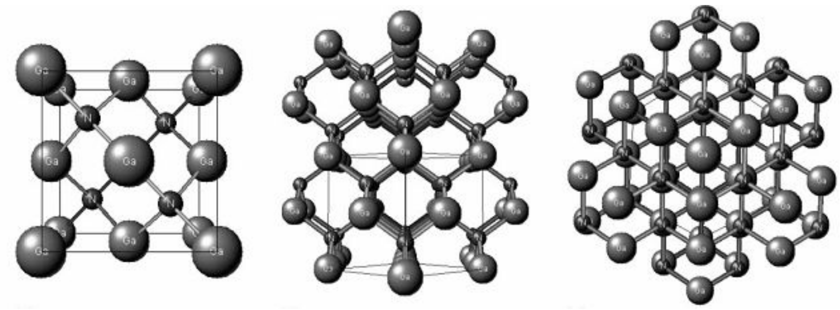
2.1 Structure of GaN

In stable condition Gallium Nitride has Wurtzite structure [3]. It has alternating biatomic close-packed (0001) planes of Ga and N pairs stacked in an ABABAB sequence. Fig. 1 shows the [0001], [$11\bar{2}0$], and [$10\bar{1}0$] directions, among which [0001] is closed-packed. The most favorable orientation for growth of smooth GaN film is [0001]. There is another structure of GaN, called the Zincblende structure. This structure of GaN (shown in Fig. 2), having the (111) closed-packed planes can be stabilized in the epitaxy.



a b c

Fig. 1 *– Perspective views of wurtzite GaN along various directions:* [0001] *(a);* [$11\bar{2}0$] *(b);* [$10\bar{1}0$] *(c) (Ref. [5])*



a b c

Fig. 2 *– Perspective views of Zincblende GaN along various direction:* [1000] *(1 × 1 × 1 unit) (a);* [110] *(2 × 2 × 2 units) (b);* [111] *(2 × 2 × 2 units) (c) [5]*

to Contents, page 1

2.1 Dislocation Densities of GaN

Generally for the heteroepitaxial growth of GaN, two types of dislocations are found: (i) Misfit Dislocation (MD) and Threading Dislocation (TD). Generally, there are three kinds of TDs [5-7]; a pure-edge dislocation with Burgers vector $\vec{b} = 1/3[11\bar{2}0]$, a screw dislocation with $\vec{b} = [0001]$, and a mixed dislocation with $\vec{b} = 1/3[11\bar{2}\bar{3}]$. The structures of the Burgers vectors in GaN film is shown in Fig. 3 [8]. According to Dadgar et al. [9], pure-edge dislocation lying in the [0001] direction, is predominant, in GaN.

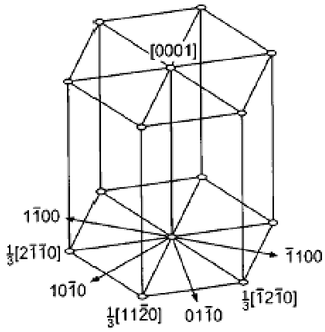


Fig. 3 *– Directions and Burgers vector notations for the hexagonal structure of GaN (Ref. [8])*

Table 1 shows the mismatch between the GaN layer with Si, SiC and Sapphire substrate. The MD and TD for heteroepitaxy of GaN layer grown on sapphire and SiC substrate is found to be of the order of 10^8 - 10^{10} cm^{-2} while that for homoepitaxy on Si substrate is 10^2 - 10^4 cm^{-2} [10]. Other defects like stacking faults, inversion domain boundaries are also obtained during heteroepitaxy [11]. These also affect the GaN-made semiconductor device quality, by reducing charge carrier mobility, minority carrier lifetime, and thermal conductivity.

Table 1 *– Lattice constants and lattice mismatches of GaN with various substrates*

Material	Structure	Lattice Constants			Lattice Mismatch (In %)
		a	*b*	*c*	
W-GaN	Wurtzite	0.31885	-	0.5185	NA
SiC	4H-W	0.3073	-	1.0053	~ 3.1 %
Si	Diamond	0.5431	-		~ 16.9 %
Sapphire (Al_2O_3)	Rhombohedra*l*, Hexagonal	0.4765	-	1.2982	~ 15 %

3. GAN FILM ON SAPPHIRE SUBSTRATE

3.1 Choice of plane of sapphire substrate for growth of GaN

Due to the large lattice-mismatch between sapphire and GaN, many threading dislocations (~ 10^{10} cm^{-2}) are found along the c-plane of the interface of the epitaxial layer of GaN grown on sapphire substrate.

to Contents, page 1

Coalescence of the nucleation islands are generated by the thermal treatment of buffer layers before the growth of high temperature GaN over-layers. It creates low angle grain boundary, resulting in threading dislocations along the c-axis. There are other planes of sapphire, which could be chosen for growth of GaN. Specially, a-plane of sapphire can be easily cleaved along r-plane as well as oriented along [0001] direction, resulting in lower lattice mismatch (2 % only) [3]. But c-plane is extremely important for growing GaN. The Al-O bond in the sapphire surface is anti-parallel to the N-III bonds in the III-N films. As a result of that AlN nucleation layer is formed by the growth technique and it is asymmetrically strained for *a*-plane of sapphire [12]. The *r*-plane and m-plane of sapphire are very rough, since inverted twins are formed there creating very high dislocation density [3]. Thus, analyzing the PL and HRXRD spectra, it is concluded that c-plane of sapphire is the best for growing GaN.

3.2 Disadvantages of sapphire substrate

Large lattice mismatch between sapphire and GaN gives a complete relaxation (no strain) for the GaN film during growth, though it increases dislocation density. Another important aspect is availability of sapphire substrate, making the process is very cheap. Moreover, sapphire is electrically insulator. Hence, the available area for device operation decreases, but it creates a low leakage of current [3].

3.3 Advantages of sapphire substrate

Lattice mismatch of sapphire with GaN is very high (~ 15%) [3]. It creates a very large dislocation density (1010 cm^{-2}) [3], which creates (a) Low minority carrier lifetime, (b) Low charge carrier mobility, (c) Low thermal conductivity. All of these aspects reduce the efficiency of the GaN-made device. Sapphire has higher thermal expansion co-efficient than GaN, which creates biaxial stress, generating crack [13]. Cleavage plane of epitaxial GaN is not parallel to those of sapphire, for which facet formation is difficult [3].

3.4 Reduction of threading dislocation density of GaN on sapphire substrate

Growth of GaN on planar substrate of sapphire, SiC or Si (111) substrate gives high TD (~ 10^{10} cm^{-2}) density. Longevity of p-n junction devices, GaN-made sophisticated devices like LEDs, decreases and leakage current increases to a very high value. This can be solved only when TD density decreases below 10^6 cm^{-2}.

To minimize the threading dislocation density, nucleation layer growth on c-plane sapphire substrate for different conditions were studied by S. Keller et al. [1]. They showed that exposure of GaN film to NH_3 before deposition of GaN layer by MOCVD reduces dislocation density. The dislocation density reduces from 2×10^{10} to 4×10^8 cm^{-2} for shorter NH_3 preflow times and for symmetric plane [1].

Kurai et al. has made high-thickness-low-diameter epitaxial GaN layer on sapphire substrate using sublimation method [15]. Now let us consider the case of growth of GaN film selectively, using HVPE (Hydride Vapor Phase Epitaxy). Several-hundred-micron-thick high quality GaN with ZnO buffer layer was selectively grown by HVPE using GaCl and NH_3 [16-18] at high

to Contents, page 1

growth rate [19], but cracks occur in the layer [20]. To overcome this problem, using MOVPE (Metalorganic vapor phase epitaxy), a thin layer (thickness 1 - 1.5 μm) of GaN was grown on the sapphire substrate first, then very thick layer (up to 4 μm) of crack-free GaN (growth temperature 1000 °C) was grown [21]. The schematic diagram of the structure is shown in Fig. 4.

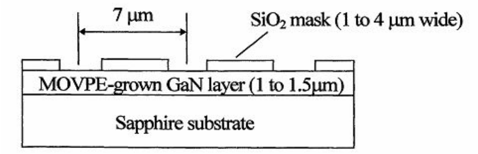


***Fig. 4** – Schematic diagram of the substrate structure (Ref. [21])*

Moreover, the coalescence of the selectively grown structure in this process abolishes the gap between the facets and finally makes a [0001] mirror-like flat surface. The defect density was measured by TEM, which gave a very low value (~ 6×10^7 cm^{-2}). Dislocations provide very low resistance against strain due to thermal expansion co-efficient difference, making a harder GaN film.

The same technique was applied by Sakai et al. to reduce the dislocation density. Selective growth of GaN layer by HVPE, in which growth on SiO_2-striped-patterned GaN layers was made by MOVPE on sapphire [0001] substrates (Fig. 5) [8].

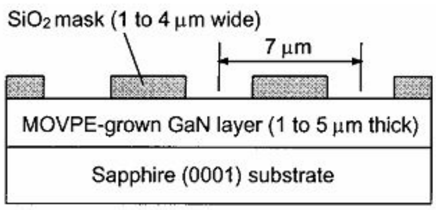


***Fig. 5** – Schematic diagram of the substrate structure used for HVPE growth (Ref. [8])*

When TEM was performed for both the layers (i.e., MOVPE grown and HVPE grown), it was observed that for MOVPE grown layer, most of the threading dislocations were pure-edge threading dislocations, some are of mixed character and very few are screw-type [5-7], which were vertically aligned. Almost all the dislocations in MOVPE layer propagated into HVPE layer laterally during the selective growth of the HVPE layer. The dislocations in HVPE layer parallel to the interface formed an angled configuration, generating no defect in the interface. These angled dislocations piled up with the [0001] direction. TEM reveals that angled dislocation depends on their Burger vectors of either $1/3[11\overline{2}\overline{0}]$ or $1/3[11\overline{2}\overline{3}]$. Since the pure-edge threading dislocations changed into screw dislocation after lateral propagation from MOVPE layer to HVPE layer, it was concluded that most of the dislocation cannot thread the HVPE layer [8]. The lateral propagation occurs mainly around the SiO_2 mask for non-pure-edge dislocations (which are very few in MOVPE layer). Moreover the lateral segments do not lie on the slip planes. This kind of morphology reduces the threading dislocation in the thicker HVPE layer.

to Contents, page 1

A great reduction of dislocation density is obtained by Lateral Overgrowth from Trenches (LOFT) in which trenches are formed by etching the GaN layer (with threading dislocation 8×10^9 cm^{-2}) grown on sapphire substrate where the GaN layer is regrown (2 - 6 μm with reduced threading dislocation 6×10^7 cm^{-2}) laterally [22]. The residual dislocations are mainly due to the merging between two lateral growths or extended from the trench sidewalls. Schematic diagram of the structure is shown in Fig 6.

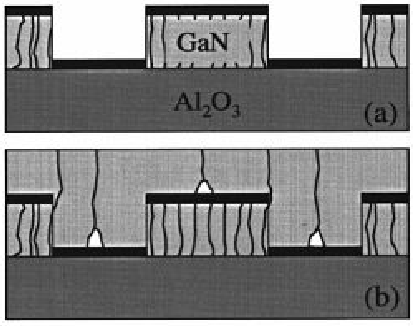


Fig. 6 *– The structure of GaN film before and after regrowth are shown schematically in (a) and (b) respectively. The TD and the voids possibly formed in GaN thin films are also shown schematically. (Ref. [22])*

A new method of applying NH_3 and SiH_4 gases simultaneously at low temperature with a certain time before the growth of a low-temperature (450 °C) GaN buffer layer followed by growth of undoped GaN layer (growth temperature 1075 °C and thickness 2μm) by MOCVD decreases the threading dislocation from 7×10^8 cm^{-2} to almost invisible under TEM, observed by Wang et al. [23]. Introduction of SiH_4 and NH_3 creates nano-sized holes, which increases the lateral growth to the next layer creating reduction of TDs. In practice, this can also be done by deposition of SiN layer (2 nm, deposition time 125 s) on sapphire, which covers the substrate up to 90 % to make 66 % of the surface area almost dislocation-free [23].

In undoped GaN increasing staking faults plays a crucial role in the reduction of dislocation density [24]. GaN buffer layers of various thicknesses on sapphire substrate were operated with various (from 10 - 80 μmole/min) TMGa (Trimethyle Gallium) flow-rates (f_{TMGa}). H. Cho et al. showed that the maximum TMGa flow rate (80 μmole/min) gives maximum intense PL peak related to increase in stacking faults as well as decrease in dislocations (1×10^8 cm^{-2}), as threading dislocations interact with stacking faults and bend towards [$1\bar{1}01$] planes and finally disappears [25].

Many techniques have been followed (i.e., LEO [26, 27], pendeoepitaxy [28], LOFT [22] to lower TD density below 10^6 cm^{-2}. All are time-consuming, complex, multi-step processes which need *ex-situ* lithography. A simple process for it is Cantilever Epitaxy (CE) in which pre-patterned (with narrow lines) sapphire substrate is used to provide reduced-dimension mesa regions for nucleation. Then etched trenches are employed for suspended lateral growth of GaN/AlGaN. The substrate is etched to a depth that allows coalescence of laterally growing GaN nucleation on the mesa surfaces before

to Contents, page 1

vertical growth fills the etched trench. Low dislocation density is obtained in the cantilever region, almost 1 μm less than the mesa region [29]. The SEM micrograph and CL image are shown in Fig. 8 and 9 respectively.

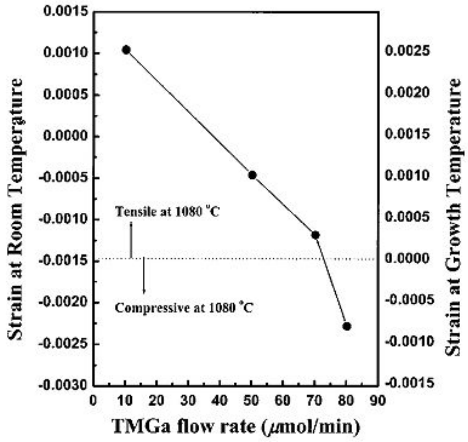


Fig. 7 – Relationship between the f_{TMGa} of buffer layers and the strain of GaN overlayers at room temperature after cool down and growth temperature (1080 °C) (Ref. [25])

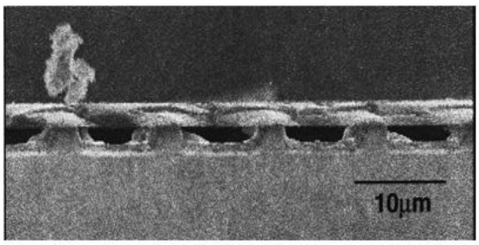


Fig. 8 – Cross-section SEM micrograph of cantilever epitaxy (Ref. [29])

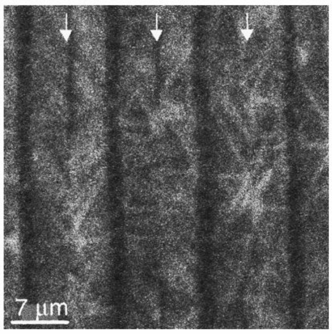


Fig. 9 – CL image of CE reducing partial absence of TDs (dark regions) along coalescence fronts (arrows). (Ref. [29])

In OMVPE process, a single thin interlayer of low temperature (700 - 900 °C for 15 minutes) can be introduced after initial growth on sapphire substrate with low-temperature buffer layer and before final growth (both at high temperature i.e., 1000 °C) to reduce the threading dislocation density (below $8 \times 10^7\ cm^{-2}$). The process is called single intermediate temperature interlayer (IT-IL), schematically shown in Fig. 10 [30]. After the sudden drop of the temperature, 3D growth mode is generated instead of 2D growth mode (generated in the initial layer) and after increase in temperature again 2D growth mode is generated again. Large amount of threading dislocations present into initial layer are bent into the interlayer, become confined as

to Contents, page 1

these cannot propagate in the 2D mode from 3D and are not exposed in the final surface [30].

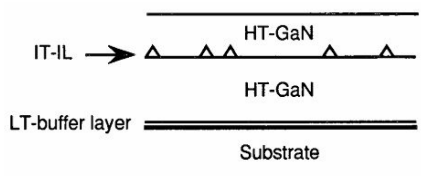


Fig 10 – Schematic diagram of IT-IL process. Ref. [30]

Another important technique to grow GaN layer with lower dislocation density (10^7 cm^{-2}) is by growth interruption modulation in HVPE method with modulating the growth process by switching on/off the GaN layer. In this process, thick multilayered structure of GaN is grown on c-Al_2O_3 (i.e., c-plane of sapphire) substrate in which dislocation reduces by and by from the lowermost layer (10^{10} cm^{-2}) towards the uppermost layer (10^7 cm^{-2}). During growth, continuous flow of NH_3 and periodic flow of HCl (shown in figure 11) are applied to modulate the growth [31].

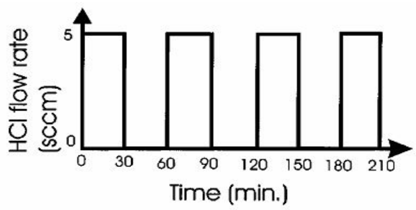


Fig. 11 – Time chart of HCl flow in GIM with deposition of 30 *min and HCl interruption of* 30 *min for each run. Ref. [31]*

K. Pakula et al. [32] showed that lateral overgrowth of GaN on sapphire (using MOCVD) with a growth interruption followed by annealing with SiH_4 decreases the TD density up to 5×10^7 cm^{-2}. This is due to pyramidal pits (approx. 40 nm deep) of the GaN surface, which are selectively etched by SiH_4 (Fig. 12) resulting in a radical change in direction of propagation of dislocation (being horizontal from the direction parallel to the c-plane). In contrast with other epitaxial methods, in this method large area of low dislocation density can be formed without formation of sub-grains with significant tilt [33, 34].

Shen et al. [35] proposed a method of reduction of both tilting and twisting of grain features causing high TD density of GaN grown on vicinal sapphire (0001) surface by RF-MBE making the vicinal angle larger than 0.50. Use of SiN, Si_xN_y, $Si_xAl_{1-x}N$, Si irradiation etc need complicated growth processes [23, 36, 37]. Fig. 13 shows the XRD rocking curves of the GaN surface. For better morphological GaN surface on vicinal substrate, MBE has higher advantages upon MOCVD [35] by reducing the density of TD to ~ 10^7 cm^{-2}.

to Contents, page 1

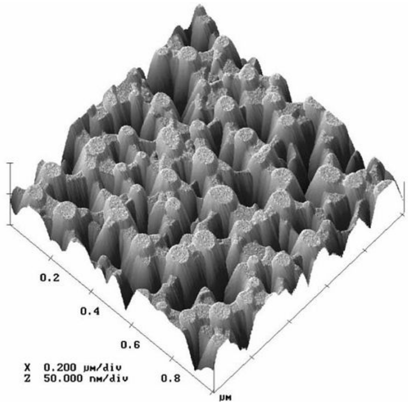


***Fig. 12** – AFM image of GaN surface after* 30 *s in situ treatment by silane at* 1100 *°C (Ref. [32])*

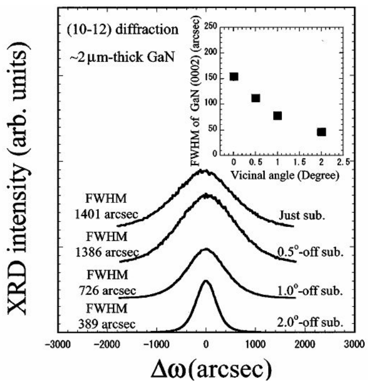


***Fig. 13** – Asymmetric (*10-12*) rocking curve of GaN films grown on vicinal sapphire* [0001] *substrate with various vicinal angles. The inset is the dependence of the FWHM of the symmetric* [0002] *diffraction peaks on the vicinal angles. Ref. [35]*

Use of in-situ thin discontinuous SiNx interlayer (deposited at 860 °C) coverage for GaN growth on c-plane sapphire (0001) substrate by OMVPE can reduce TD density up to a factor of 50 to 9×10^7 cm^{-2} mainly due to construction of facetted islands on the GaN surface as well as generation of half-loops between bent-over TDs during the lateral overgrowth [38]. Dependence of TD density with the interlayer thickness is shown in Fig. 14. Here c-plane sapphire with 0.25 % miscut towards the a-axis is used as substrate. Same kind of growth by MOCVD shows TD density to be 4.4×10^7 (screw-type) and 1.7×10^7(edge-type) [39].

Chakraborty et al. [40] showed that SiN_x *in-situ* nanomask is also important for TD density reduction in non-polar a-plane GaN films on r-plane sapphire by MOCVD. Increase in SiN_x layer leads to on-axis and off-axis FWHM of HRXRD, rms surface roughness and submicron pit density as well as stacking fault density decrease from 8.0×10^5 to 3.0×10^5 cm^{-2} which finally decreased defects (both SFs and TDs) up to 9×10^9cm^{-2}. Increase in deposition time leads to decrease in TDs.

to Contents, page 1

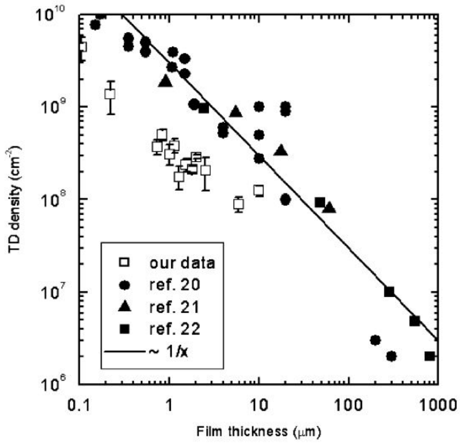


***Fig. 14** – TD density vs. GaN layer thickness of SiN_x compared with HVPE-grown sample from Mathis et al. (Ref. [20]), Morkoc (Ref. [21]) and Lee et al. (Ref. [22]). The solid line represents the trend of the TD reduction with thickness of the HVPE-grown layers. (Ref. [39])*

C.J. Tun et al. showed that multiple Mg_xN_y/GaN buffer layers, which can form a textured surface, were used for growth of GaN on sapphire using MOCVD at 530 °C. Reduction of dislocation (associated with the grain boundary) was obtained. Propagation of TD after coalescence occurs here, lowering its density on the exposed surface, as shown in Fig. 15 [41].

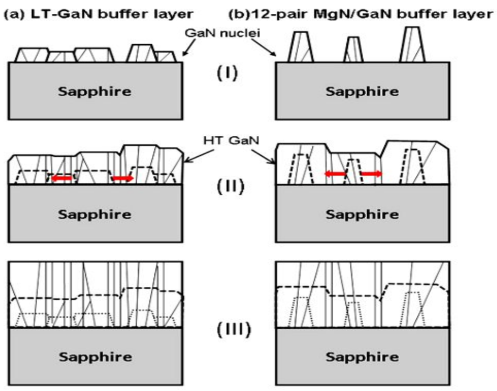


***Fig. 15** – (color online) schematic representations of the morphological evolution and associated TDs generation and propagation in GaN epitaxial layers grown on (a) LT-GaN, (b) 12 pairs of Mg_xN_y buffer layers on sapphire: (I) heat-treated buffer layers on sapphire, (II) coalescence of HT-GaN with accompanying TDs generation, (III) propagation of TDs after coalescence. (Ref. [41])*

M.A. Moram et al. [42] showed that a 500 nm deposition of lattice-matched, dislocation-blocking scandium nitride interlayer introduction using a single step without lithography for growth of GaN on sapphire template can reduce TD density reduction up to ~ $10^7 cm^{-2}$ for coalesced films and ~10^6 cm^{-2} for partially-coalesced film. Reduction of TD density depends on the thickness of scandium interlayer, as shown in Fig. 16. The reasons of TD

to Contents, page 1

density reduction are – (1) limited chemical stability under MOVPE growth condition, (2) void formation and bending of dislocation during annealing, creating no threading of dislocations in exposed template area, (3) matching of interatomic spacing between GaN and ScN mask layer, resulting in nucleation of overgrown GaN islands without forming dislocation.

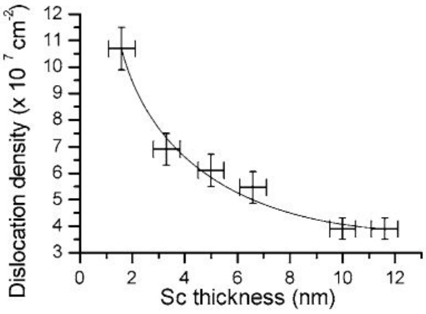


Fig. 16 *– Plot of TD density of coalesced GaN layer vs. Scandium thickness (Ref. [42]*

Q. Li et al. used self-assembled close-packed monolayer of silica microspheres as selective growth mask for growth of GaN on sapphire epilayer. Silica microspheres are formed during regrowth of GaN layer, which terminated the propagation of dislocation, causing a huge reduction of dislocation by bending and blocking (~ 4×10^7 cm^{-2}) [43]. Fig. 17 shows the dependence of TD density with the silica microsphere diameter.

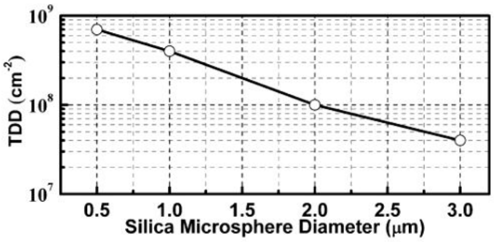


Fig. 17 *– TDD of GaN layer as a function of silica sphere diameter (Ref. [43])*

Dislocation density of three-fourth of the exposed area of GaN, grown on sapphire substrate was reduced extremely (~ 10^2 cm^{-2}) during the two-step method of LPE growth using Na-flux [2]. The thickness of the GaN crystal was made intentionally very high (2 mm) here to show that dislocation reduces with increase in growth thickness. The schematic diagram of the mechanism of TD reduction in this process is given in Fig. 18.

4. GAN FILM ON SIC SUBSTRATE

Relatively high quality of GaN epitaxial device on Sapphire substrate is a bit difficult to achieve because of their large lattice mismatch (~ 15 %). On the other hand better lattice mismatch between GaN and Silicon Carbide as a substrate has gained popularity in recent years for both MOCVD/MOVPE and MBE growth technology.

to Contents, page 1

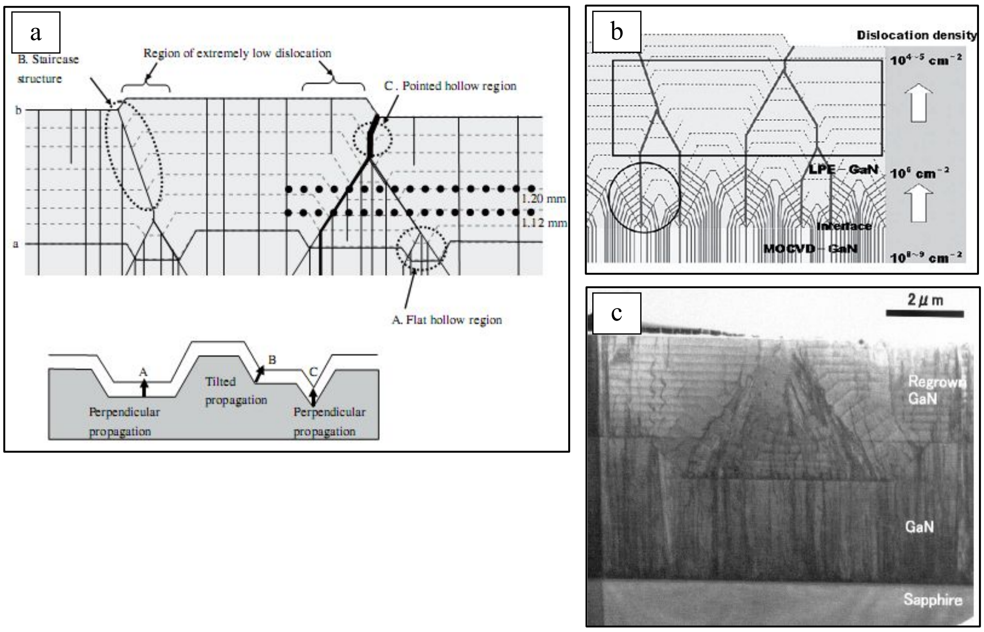


Fig. 18 *– The dislocation reduction process after the middle stage of the LPE growth. Dislocations existing at positions A and C propagate along the c-axis, while those at the staircase structures (positions B) propagate along the development of the stair edge, resulting in oblique propagation from the c-axis. By following the bold line in the figure, we can summarize the propagation of the dislocations as follows: Type A dislocations are aggregated to Type B ones as the LPE growth progresses, followed by their aggregation to Type C and eventually back to Type B. Although the transformation of the dislocations from Type B to Type C and subsequently back from Type C to Type B occurred naturally, transformations into Type A were not allowed after extinction of the concave portions. Therefore, the dislocation types transformed in the order A, B, C, B, C, B ..., reducing the number of dislocations. Ultimately, almost all dislocations were categorized as Type B (a). The entire process of dislocation reduction, including the Initial growth stage, is summarized (Ref. [2]) (b). TEM shows very low DD in the grooved side-walls [59] (c)*

The process which creates lowest DD is ELOG (epitaxial lateral overgrowth). Here, regrowth technique is applied on the periodically grooved surface with controlled V/III ratio. Groove is completely buried in the thinner exposed layer, for which DD is unexposed in TEM. It creates very low DD (6×10^6 cm^{-2}) in the regrown GaN surface. Fig. 18c shows the TEM of this phenomenon [59].

to Contents, page 1

4.1 Comparison of SiC substrate over sapphire for growth of GaN layer

Table 2 *– Advantages of SiC in comparison with sapphire*

Parameters	SiC	Sapphire
Lattice constant mismatch	3.1 % along [0001] direction, but contribution of Screw Dislocation creates large overall DD	Very large (approx. 15 %), creating high DD.
Thermal conductivity[11]	3.8 W/cm K	0.25 W/cm K (<< SiC)
Electrical conductivity	Relatively higher value	Insulator
Orientation of crystal planes with respect to GaN	Parallel (facet formation easier by cleaving)	Not parallel (facet formation is not easy)
Polarity	GaN film polarity is easier on SiC substrate, as it has both C and Si polarity.	GaN film polarity is problematic on sapphire, as mentioned before in this review.
Surface roughness	Higher than sapphire (1 nm rms), disadvantageous with respect to sapphire.	Lower than SiC (0.1 nm rms), better for GaN growth.
Thermal expansion coefficient	Less than GaN, generating bi-axial tension.	Higher than GaN, generating bi-axial stress.

4.2 Choice of SiC plane as substrate

SiC has 250 polytypes, among which only two (4H-SiC and 6H-SiC) have same space group {p6_3mc (no. 186)} as Wurtzite GaN. 6H-SiC is more commercially available than 4H-SiC. Thus, 6H-SiC is the mostly used polytypes of GaN growth on SiC substrate [3].

4.3 Reduction of threading dislocation density of GaN film on SiC substrate

The main problematic TD in SiC substrate is hollow-core screw dislocations (called micropipes or nanopipes), which have Burgers vector ($\vec{b}$) twice and three times of c-lattice constant of 6H-SiC and 4H-SiC respectively. These holes are oriented along the c-axis; as a result, these can propagate throughout the crystal [45]. Here are some techniques of reduction of dislocation density of GaN layer grown on SiC substrate.

The lateral overgrowth via Organometalic Vapor Phase epitaxy (OMVPE) of GaN stripes patterned in a SiO_2 mask deposited on GaN-film/AlN-buffer/6H-SiC (0001)-substrate at 1000 - 1100 °C and at 45 Torr, oriented along $[11\bar{2}0]$ and $[1\bar{1}00]$ is required here. Fig. 19 shows the schematic diagram. A very low dislocation density of 10^6 cm^{-2} GaN layer is obtained due to coalescence of the GaN homoepitaxial stripes on SiO_2 mask. The coalesced layers for this case had a surface roughness of rms value, i.e., 0.25 nm [46].

The low-edge grain boundary during growth of GaN on SiC substrate is the main source of dislocation for this heterostructure. Two possible sources of dislocation are (i) dislocation half-loop and (ii) island edges. Accumulation of misfit strains after a certain thickness creates misfit, which ultimately generates TDs. The GaN layer (1.5 μm)/AlN buffer/SiC (typically tilted 3 to 4 degrees off towards the direction $[11\overline{20})$ substrate grown by MOCVD using ultra-thin AlN buffer layer of 1.5 nm (< the critical value for misfit dislocation) and a smooth AlN surface reduces the threading dislocation very much (2-3 order less than the value of dislocation for the thicker buffer). Island edges can be reduced by smooth surface. Hence TD density reduction is obtained [5].

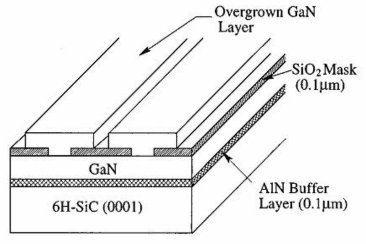


Fig. 19 *– Schematic diagram showing lateral epitaxial overgrowth of GaN layer on SiO2 mask from GaN deposited within striped window openings on GaN/AlN/6H–SiC substrates (Ref. [46])*

The defects in the GaN grown within the SiO_2 windows were predominantly threading dislocations of mixed character with Burgers vector $b = 1/3[11\bar{2}3]$ and edge dislocations with $b = 1/3[11\bar{2}0]$. Hexagonal pyramids of GaN are grown under the SiO_2 mask during lateral epitaxial overgrowth (LEO) of GaN on patterned GaN/AlN/6H-SiC substrate (Fig. 20a). Fig. 20b shows the schematic diagram of LEO. Stresses due to the mismatches in the thermal expansion coefficients during LEO-GaN growth, existing initially at the underlying AlN/6H-SiC and GaN/AlN interfaces, propagate through the selectively grown GaN and are accommodated in these regions both via the continued propagation of numerous TDs and short dislocation segments and the bending of the stripes or pyramids. As a result, LEO-GaN has lower dislocation density (approximately 4 order less) than the vertical-growth region [47, 48].

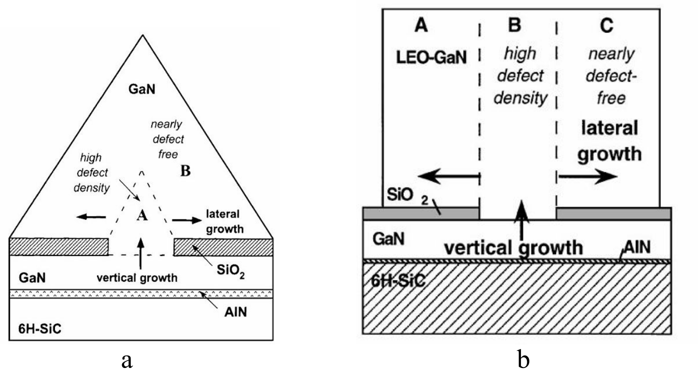


Fig. 20 *– Lateral Epitaxial Overgrowth of selectively grown GaN hexagonal pyramid (a). A schematic diagram of the lateral epitaxial overgrowth (LEO) in a selectively grown GaN stripe (b) (Ref. [48])*

Selectively grown (i.e., using SiO_2 mask) GaN hexagonal pyramids and stripes on circularly patterned GaN [0001] on sapphire [49-51] and 6H-SiC

to Contents, page 1

[0001] substrate [52, 53], i.e., on GaN/AlN/6H-SiC heterostructure by the lateral epitaxy method (a two-stage mechanism containing vertical and lateral growth [54]) has also low dislocation density [27]. Nearly defect-free single crystal GaN is obtained during the lateral growth while analyzing with TEM. The thicknesses of the AlN buffer layer and the GaN layer are 1000 Å and 1.75 μm. The coalescence of the laterally grown volumes yields nearly defect-free regions. Curved surfaces are formed within the heterostructure, to accommodate mismatches due to the co-efficient of thermal expansions among the different phases [48].

GaN film grown on SiC by plasma-assisted MBE can reduce dislocation density significantly. Reduction of width of TEM $[10\bar{1}2]$ rocking curves with reduction of Ga/N flux ratio results in the change in morphology from flat to rough, creating reduction in edge-dislocation by cluster formation with topological valleys of the rough surface [55].

5. GAN ON SILICON SUBSTRATE

5.1 Advantages of Si substrate

Very low price is the most important advantage for choosing it as substrate. It is available in very large size; it has good thermal stability too. These aspects are suitable for GaN growth. Moreover, crystal perfection is very high in comparison with the other substrates for growth of GaN.

5.2 Disadvantages of Si Substrate

Lattice constant mismatch with GaN is very high (~ 16.9 % [56]). Thermal expansion co-efficient mismatch with GaN is also very high. Moreover, tendency of formation of amorphous SiN is noticeable during growth of GaN on it. All of these aspects are responsible for high defective interface formation (between GaN layer and Si substrate).

5.3 Choice of orientation for Si substrate for GaN growth

Generally (111) Si is taken as substrate for GaN growth. The GaN-made device quality using this substrate has much improved quality than any other orientation. It can support 2-D growth of GaN.

5.4 Reduction of Threading Dislocation density for GaN layer grown on Si substrate

Dadgar et al. showed that buffer layers were used to reduce TD density of GaN layer grown on Si substrate. The large mismatch (17 %) between GaN (a = 3.1891 Å) and Si (111) (a = 3.8403 Å) generated a biaxial tensile stress in the heterostructure [57]. Misfits and cracks were generated due to residual tensile stress [57]. Reduction of tensile stress as well as dislocation density was achieved by partial masking. For *in situ* masking of AlN seed layer with thin SiN mask [9] GaN-layer quality is improved. A thin SiN mask was used here. Fig. 22 shows that reduction of stress depends on the deposition time of SiN layer.

Insertion of Si_xN_y layer (along with AlN buffer layer grown by MOCVD) for GaN grown on a Si (111) layer, as shown in Fig. 23, reduced dislocation density much more. The growth temperature was 1080 °C here. With the increasing growth time, rms surface roughness decreases, pits due to TD decreases as Si_xN_y layer acted as a filter for TD [57].

to Contents, page 1

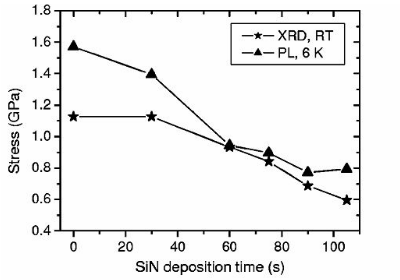


Fig. 22 *– Stress of the samples determined from PL (triangles) and x-ray diffraction (stars) measurements vs SiN deposition time. (Ref. [9])*

Growth time > 100 s can remove all types of dislocations (misfit dislocations too) resulting in a high quality GaN layer. K. Cheng et al. showed that instead of MOCVD, if high temperature MOVPE is applied for the same structure using a combination of AlGaN intermediate layer along with Si_xN_y layer, smooth and fully coalesced layer of dislocation density 3.0×10^8 - $5.0 \times 10^{8-2}$ cm can be obtained [58].

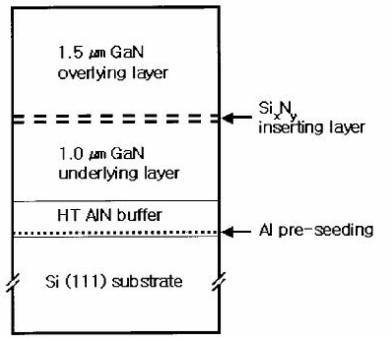


Fig. 23 *– Schematic cross-section of GaN epilayer (Ref. [57])*

6. CONCLUSION

Here we have studied different techniques of reduction of dislocation density of GaN layer grown on foreign substrates. We have chosen here three most commonly used substrates: sapphire, SiC and Si. We have studied the lattice mismatching percentage at the interface between the substrate and layer as well as other reasons of threading dislocations. reduction of dislocation density during the different growth techniques like MOCVD, MBE, HVPE, MOVPE etc., on three above mentioned substrates have been studied here separately.

Current researchers prefer sapphire substrate when comparing with SiC, though it has much higher lattice mismatch, thermal conductivity, difficulty in facet formation etc. Sapphire costs lower, provides relaxation to the growth and has lower surface roughness in comparison with SiC. Due to these reasons, researchers tried to reduce its dislocation density to increase its use as substrate.

Silicon has a very large lattice mismatch with GaN, resulting in very high dislocation density. Due to very low cost and thermal stability people have started working on GaN/Si growth technique recently. Still it has not extended in the industrial level. A good-quality low-cost GaN-on-Si substrate is still a good area of research.

to Contents, page 1

REFERENCE

1. D.F. Brown, R. Chu, S. Keller, S.P. DenBaars, U.K. Mishra, *Appl. Phys. Lett.* **68,** 1525 (2009).
2. F. Kawamura, M. Tanpo, N. Miyoshi, M. Imade, M. Yoshimura, Y. Moria, Y. Kitaoka, T. Sasaki, *J. Cryst. Growth* **311,** 3019 (2009).
3. L. Liu, J. Edgar, *Mat. Sci. Eng. R* 37, 61 (2002).
4. S. Porowski, J.M. Baranowski, M. Leszczynski, J. Jun, M. Bockowski, I. Grzegory, S. Krukowski, M. Wroblewski, B. Lucznik, G. Nowak, K. Pakula, A. Wysmolek, K.P. Korona, R.P. Stepniewoski. *Proc. Int. Symp. Blue Laser and Light Emitting Diodes* (Ohmsha: Tokyo: 1966).
5. W. Qian, M. Skowronski, M. De Graef, K. Doverspike, L.B. Rowland, D.K. Gaskil, *Appl. Phys. Lett.* **66,** 1252 (1995).
6. X.J. Ning, F.R. Chiena, P. Pirouza, J.W. Yanga, M. Asif Khan, *J. Mater. Res.* **11,** 580 (1996).
7. X.H. Wu, L.M. Brown, D. Kapolnek, S. Keller, B. Keller, S.P. Den Baars, J.S. Speck, *J. Appl. Phys.* **80,** 3228 (1996).
8. A. Sakai, H. Sunakawa, A. Usui, *Appl. Phys. Lett.* **71,** 2259 (1997).
9. A. Dadgar M. Poschenrieder, A. Reiher, J. Blasing, J. Christen, A. Krtschil, T. Finger, T. Hempel, A. Diez, A. Krost, *Appl. Phys. Lett.* **82,** 28 (2003).
10. B. Gil, Group three nitride semiconductor compounds (Clarendon Press: 1998).
11. J. Edgar, Properties, processing and applications of gallium nitride and related semiconductors (INSPEC: 1999).
12. P. Vennegues, Z. Bougrioua, *Appl. Phys. Lett.* **89,** 111915 (2006).
13. E. Etzkorn, D. Clarke, *J. Appl. Phys.* **89,** 1025 (2001).
14. J. Van Nostrand J. Solomon, A. Saxler, Q.-H. Xie, D.C. Reynolds, D.C. Look, *J. Appl. Phys.* **87,** 8766 (2000).
15. S. Kurai,Y. Naoi, T. Abe, S. Ohmi, S. Sakai, *Jpn. J. Appl. Phys. 2* **35,** 77 (1996).
16. K. Naniwae, Sh. Itoh, H. Amano, K. Itoh, K. Hiramatsu and Is. Akasaki, *J. Cryst. Growth* **99,** 381 (1990).
17. W. Gotz, L.T. Romano, B.S. Krusor, N.M. Johnson, R.J. Molnar, *Appl. Phys. Lett.* **69,** 242 (1996).
18. Y. Melnik, p. 863 (Bristol (England); Boston: Adam Hilger Ltd.: 1996).
19. T. Detchprohm, K. Hiramatsu, H. Amano, I. Akasaki, *Appl. Phys. Lett.* **61,** 2688 (1992).
20. K. Hiramatsu, T. Detchprohm, I. Akasaki, *Jpn J. Appl. Phys.* **32,** 1528 (1993).
21. J. Westwater, D.P. Gosain, S. Usui, *Jpn. J. Appl. Phys.* **2,** 6204 (1997).
22. Y. Chen, R. Schneider, S.Y. Wang, R.S. Kern, C.H. Chen, C.P. Kuo, *Appl. Phys. Lett.* **75,** 2062 (1999).
23. T. Wang, Y. Morishima, N. Naoi, S. Sakai, *J. Cryst. Growth* **213,** 188 (2000).
24. S.I. Molina, A.M. Sanchez, F.J. Pacheco, R. Garcia, M.A. Sanchez-Garcia, F.J. Sanchez, E. Calleja, *Appl. Phys. Lett.* **74,** 3362 (1999).
25. H. Cho J. Y. Lee, Ki Soo Kim, Gye Mo Yang, Jae Ho Song, and Phil Won Yu, *J. Appl. Phys.* **89,** 2617 (2001).
26. D. Kapolnek, S. Keller, R. Vetury, R.D. Underwood, P. Kozodoy, S.P. Den Baars, and U.K. Mishra, *Appl. Phys. Lett.* **71,** 1204 (1997).
27. T.S. Zheleva, O.-H. Nam, M.D. Bremser, R.F. Davis, *Appl. Phys. Lett.* **71,** 2472 (1997).
28. K. Linthicum, Th. Gehrke, D. Thomson, E. Carlson, P. Rajagopal, T. Smith, D. Batchelor, R. Davis, *Appl. Phys. Lett.* **75,** 196 (1999).
29. I.H. Carol, C.C. Mitchell, J. Han, N.A. Missert, P.P. Provencio, D.M. Follstaedt, Gr.M. Peake, L. Griego, *Appl. Phys. Lett.* **77,** 3233 (2000).
30. E.D. Bourret-Courchesne, S. Kellermann, K.M. Yu, M. Benamara, Z. Liliental-Weber, J. Washburn, S.J.C. Irvine, A. Stafford, *Appl. Phys. Lett.* 77, 3562 (2000).

to Contents, page 1

31. W. Zhang, St Roesel, H. R. Alves, D. Meister, W. Kriegseis, D. M. Hofmann, Br. K. Meyer, Till Riemann, P. Veit, J. Blaesing, Al. Krost, and J. Christen, *Appl. Phys. Lett.* **78,** 772 (2001).
32. K. Pakula, R. Bozek, J.M. Baranowski, J. Jasinski, Z. Liliental-Weber, *J. Cryst. Growth* **267,** 1 (2004).
33. J.J.Z. Liliental-Weber, D. Cherns, M. Baines, R. Davis *Materials Research Society Proceedings* **693**, 309 (2001).
34. Z. Liliental-Weber, D. Cherns, *J. Appl. Phys.* **89,** 7833 (2001).
35. X. Shen, H. Matsuhata, H. Okumura, *Appl. Phys. Lett.* **86,** 021912 (2005).
36. X. Shen, H.-K. Cho, T. Ide, M. Shimizu, H. Okumura *Jpn. J. Appl. Phys.* **41,** 1428 (2002).
37. S. Tanaka, M. Takeuchi, Y. Aoyagi, *Jpn. J. Appl. Phys*. 2 **39,** 831 (2000).
38. M. Kappers, R. Datta, R.A. Oliver, F.D.G. Rayment, M.E. Vickers, C.J. Humphreys, *J. Cryst. Growth* **300,** 70 (2007).
39. J. Xie, U. Ozgur, Y. Fu, X. Ni, H. Morkoc, C.K. Inoki, T.S. Kuan, J.V. Foreman, H.O. Everitt, *Appl. Phys. Lett.* **90,** 041107 (2007).
40. Arp. Chakraborty, K.C. Kim, F. Wu, J.S. Speck, S.P. DenBaars, U.K. Mishra, *Appl. Phys. Lett.* 89, 041903 (2006).
41. C.J. Tun, C.H. Kuo, Y.K. Fu, C.W. Kuo, C.J. Pan, and G.C. Chi, *Appl. Phys. Lett.* **90,** 212109 (2007).
42. M. Moram Y. Zhang, M.J. Kappers, Z.H. Barber, C.J. Humphreys, *Appl. Phys. Lett.* **91,** 152101 (2007).
43. Q. Li, J. Figiel, G. Wang, *Appl. Phys. Lett.* **94,** 231105 (2009).
44. H. Lahreche, M. Leroux, M. Laugt, M. Vaille, B. Beaumont, and P. Gibart, *J. Appl. Phys.* **87,** 577 (2000).
45. M. Dudley, and X. Huang, (Trans. Tech. Publ,: 2000).
46. O. Nam, M. D. Bremser, T. S. Zheleva, and R. F. Davis, *Appl. Phys. Lett.* **71,** 2638 (1997).
47. S. Tanaka, S. Iwai, Y. Aoyagi, *J. Cryst. Growth* **170,** 329 (1997).
48. T.S. Zheleva, O.-H. Nam, W.M. Ashmawi, J.D. Griffin, R.F. Davis, *J. Cryst. Growth* **222,** 706 (2001).
49. R. Underwood, D. Kapolnek, B.P. Keller, S. Keller, S.P. Denbaars, U.K. Mishra, *Solid-State Electron.* **41,** 243 (1997).
50. Y. Kato, Sh. Kitamura, K. Hiramatsu and N. Sawaki, *J. Cryst. Growth* **144,** 133 (1994).
51. S. Kitamura, K. Hiramatsu, N. Sawaki, *Jpn. J. Appl. Phys. 2* **34,** 1184 (1995).
52. O. Nam et al. (1996) pp. 107.
53. O. Nam et al., *Growth* **1** (1997).
54. E. Givargizov, Oriented crystallization on amorphous substrates (Plenum Press: New York: 1991).
55. C. Lee, A. Sagar, R.M. Feenstra, C.K. Inoki, T.S. Kuan, W.L. Sarney, L. Salamanca-Riba, *Appl. Phys. Lett.* **79,** 3428 (2001).
56. S. Ren, J. Dow, *Appl. Phys. Lett.* **69,** 251 (1996).
57. K. Lee, E. Shin, K. Lim, *Appl. Phys. Lett.* **85,** 1502 (2004).
58. K. Cheng, M. Leys, S. Degroote, M. Germain, G. Borghs, *Appl. Phys. Lett.* **92,** 192111 (2008).
59. M. Ishida, M. Ogawa, K. Orita, O. Imafuji, M. Yuri, T. Sugino, K. Itoh, *J. Cryst. Growth* **221,** 345 (2000).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.85-91

PACS numbers: 73.61.Ga,71.23.Cq, 61.05.Cp, 68.37.Hk, 78.20.Ci

EFFECT OF pH ON THE PHYSICAL PROPERTIES OF $ZnIn_2Se_4$ THIN FILMS GROWN BY CHEMICAL BATH DEPOSITION

P. Babu, M.V. Reddy, N. Revathi, K.T.R. Reddy*

Thin Film Laboratory, Department of Physics,
Sri Venkateswara University, Tirupati – 517502, India
* E-mail: ktrkreddy@gmail.com

Recently there has been much interest on the preparation and characterization of ternary semiconducting materials, mainly $ZnIn_2Se_4$ (ZIS) due to its potential applications in various fields, particularly as a buffer layer in the fabrication of heterojuction solar cells. In the present work, thin films of ZIS have been synthesized by a simple and economic method, chemical bath deposition at different pH values that vary from 9 *to* 11. *The deposition was carried out for a fixed bath temperature (T_b) of* 90 °*C and constant reaction time of* 60 *min. Ammonia and hydrazine hydrate were used as complexing agents. The chemical and physical properties of the deposited ZIS films were analyzed using appropriate techniques. The X-ray diffraction analysis revealed that the deposited films were polycrystalline and showed (*112*) peak as the preferred orientation. Scanning electron micrographs revealed that the samples had large number of granule like particles in different sizes. The optical transmittance of these samples was found to be* > 75 % *in the visible region and the evaluated energy band gap varied from* 2.15 *eV to* 2.64 *eV with the change of pH value in the range,* 9 - 11. *The detailed study of these results were presented and discussed.*

Keywords: *$ZnIn_2Se_4$, EDAX, XRD, SEM, OPTICAL PROPERTIES.*

(Received 04 February 2011, in final form 20 March 2011)

1. INTRODUCTION

The ternary chalcopyrite semiconductors, particularly the Cu-III-VI_2 compounds have proved to be the most potential absorber materials in the fabrication of heterojunction solar cells in terms of conversion efficiency and long-term stability. Among the various combinations of ternary as well as quaternary compounds, $CuInSe_2$ (CIS) and $CuInGaSe_2$ (CIGS) proved to be the successful absorbers and electrical conversion efficiencies > 16 % and 20 % [1] respectively have been achieved at laboratory level using CdS as the buffer and ZnO as the window layers. However, the usage of CdS in photovoltaic cell complicates the in-line process and generates environmental and health problems due to the toxicity of 'Cd' precursor. Alternatively materials such as ZnS, ZnSe, $ZnIn_2Se_4$, In_xSe_y, $Zn_{1-x}Mg_xO$ and $In(OH)_3$, In_2S_3 etc. have been proposed as the substitutes for CdS in CIGS solar cells [2, 3]. Among them, $ZnIn_2Se_4$ (ZIS) is one of the best alternatives for CdS as a buffer layer. As In and Se are already present in ZIS, the problem of lattice mismatch, which is commonly observed in CIS or CIGS-based solar cells can be minimized [4]. Further, ZIS has also potential applications in various other fields such as photoelectronics [5] and electro-optical memory devices [6].

to Contents, page 1

ZIS belongs to the $A^{I}B_2^{II}X_4^{IV}$ defect chalcopyrite family and the first report on $ZnIn_2Se_4$ by Hahn et al., gave the crystallographic data. It crystallizes in the tetragonal structure with the space group of S_4^2, and having the lattice parameters $a = 0.569$ nm and $c = 1.149$ nm [7]. Growth of ZIS thin films was first reported by Konagai et al. [8] using three-source co-evaporation of the constituent elements. ZIS layers have been prepared by different chemical and physical methods [9-11]. To our knowledge, there are no reports available on the preparation of ZIS films by chemical bath deposition (CBD) method. In our previous study, we reported on the physical properties and growth mechanism of ZIS films by CBD on glass substrates formed at different bath temperature [12]. In the present study, ZIS films have been grown at different pH values for a constant bath temperature and deposition time by CBD process and the effect of pH on the physical properties are investigated.

2. EXPERIMENTAL

ZIS films were deposited on Corning 7059 glass substrates by chemical bath deposition using 0.05 M aqueous solutions of zinc chloride ($ZnCl_2$), sodium selenite (Na_2SeO_3) and indium tri-chloride ($InCl_3$) as precursors. The deposition bath was prepared by taking 15 ml of aqueous $ZnCl_2$ solution in a separate beaker to which 2 ml of 80 % hydrazine hydrate is added. This makes the solution turbid due to the formation of hydroxides and the turbidity disappears when a few drops of 25 % ammonia is added to the solution bath. 30 ml of $InCl_3$ and 60 ml of sodium selenite precursors are added to the deposition bath in the same sequence. Finally the deposition bath is made-up for a total volume of 120 ml by adding distilled water. The pH of the bath was adjusted to the required value using ammonia and ammonium chloride. The mixture was stirred well using a temperature controllable magnetic stirrer. Ultrasonically cleaned Corning 7059 glass substrates were vertically dipped in the deposition bath while stirring of the solution continues. The deposition of ZIS films is carried out at a fixed bath temperature (T_b) of 90 °C and a reaction time of 60 min by varying the pH value in the range, 9 - 11, measured using DIGITAL pH meter. All the grown films appeared thick brown in color.

The elemental composition of ZIS samples was determined using the FEL Sirion energy dispersive analysis of X-ray (EDAX). The crystallinity of ZIS films was measured using a Siefert X-ray diffractometer (model 3003 TT) with Cu-K_α radiation source ($\lambda = 1.5402$ Å) while the surface morphology was evaluated using Zeiss scanning electron microscope (SEM) (model EVO MA 15). The optical transmittance measurements were performed using Hitachi U: 3400 UV-Vis-NIR spectrophotometer.

3. RESULTS AND DISCUSSION

3.1 EDAX Analysis

Fig. 1 shows the chemical composition of ZIS films evaluated using EDAX. The EDAX spectra indicated well defined peaks corresponding to Zn, In and Se in addition to O. All the grown films showed Se deficiency irrespective of the solution pH value. ZIS films deposited at a pH of 11 had Zn: In: Se in the ratio 1:2:3. The deviation in the stoichiometry of the films from 1:2:4

to Contents, page 1

(of $ZnIn_2Se_4$) might be due to the oxygen incorporation during the growth of ZIS films because of the usage of aqueous solutions in the bath or due to the incomplete reaction of the bath mixtures.

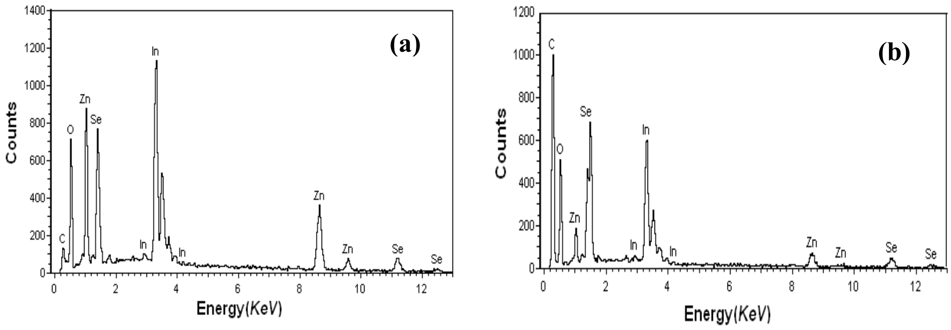


Fig. 1 *– EDAX profiles of ZIS films deposited at (a) pH = 9 and (b) pH = 11*

3.2 Structural Analysis

The X-ray diffraction (XRD) pattern of ZIS films prepared at different pH values are shown in Fig. 2. The XRD spectra of all the ZIS thin films exhibited polycrystalline nature with the tetragonal crystal structure. The diffraction pattern showed two major peaks corresponding to (110) and (112) planes of ZIS in addition to the appearance of a low intensity (204) peak. Further all the peaks were slightly shifted towards higher 2θ values with the increase of pH value. The identified planes in the XRD spectra were in good agreement with the reported data in the literature [13]. No impurity peaks such as Zn (OH), $ZnSeO_3$ etc. were found in the XRD spectra. The intensity of the (112) peak increased with the increase of pH, indicating that more grains were oriented along the (112) plane. This might be due to the formation of more hydroxides at higher pH values that act as nucleation centers and favors hydroxide cluster growth mechanism instead of ion-by-ion deposition.

The average crystallite size, D of the layers was evaluated using the Debye Scherer formula [14].

$$D = \frac{0.9\lambda}{\beta \cos\theta} \qquad (1)$$

where β is the fringe width at half maximum (FWHM) of the (112) peak, λ is the wavelength of X-rays and θ is the corresponding diffraction angle. The evaluated crystallite size varied in the range, 200 - 350 nm with the change of pH. The degree of preferred orientation corresponding to each plane was studied by the analysis of texture coefficient (C_i), which is a measure of the orientation of each reflection compared to a randomly oriented sample using the following relation [15, 16].

$$C_i = \frac{I/I_0}{(1/n)\sum_{i-1}^{n} I/I_0} \qquad (2)$$

to Contents, page 1

where I is measured intensity of the peak in the spectrum, I_0 is the intensity for completely random sample (JCPDS) and n is the number of reflections considered in the analysis. If $C = 1$, it represents the film with randomly oriented crystallites whereas $C > 1$ for films that are oriented preferentially in a particular direction. The variation of texture coefficient with pH value for (112) and (110) peaks is shown in Fig. 3. The value of $C_{(112)}$ increased linearly with the increase of pH value, whereas the value of $C_{(110)}$ was less than that of $C_{(112)}$, implying that the films were highly textured along the (112) direction.

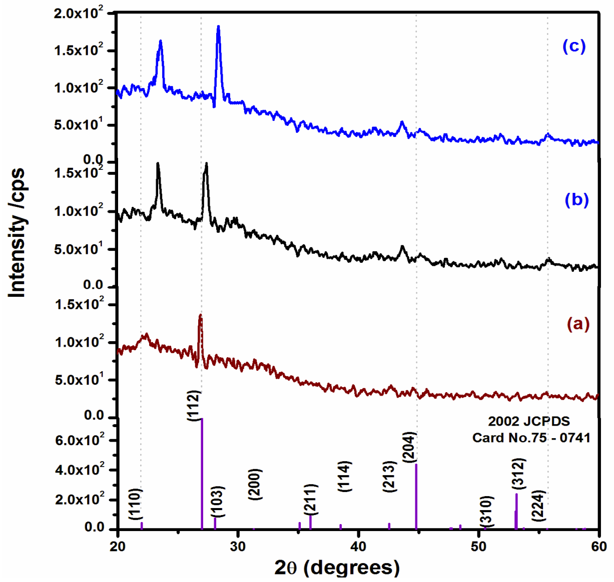


Fig. 2 – *XRD pattern of ZIS films deposited at (a) pH = 9, (b) pH = 10 and (c) pH = 11*

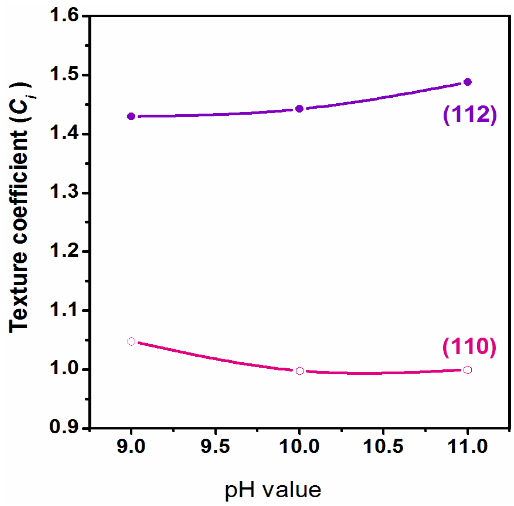


Fig. 3 – *Change of texture coefficient with pH value*

to Contents, page 1

3.3 SEM Analysis

The surface morphology of ZIS films deposited at different pH values was analyzed using the SEM images shown in Fig. 4. The micrographs revealed that the samples had large number of granule like particles with different sizes that were distributed randomly on the substrate surface. The pictures also showed groups of crystallites on an amorphous background. With the increase of pH value these groups join together forming more continuous structure compared to that of the structures grown at lower pH values. Further, the particle size also increased with the increase of pH.

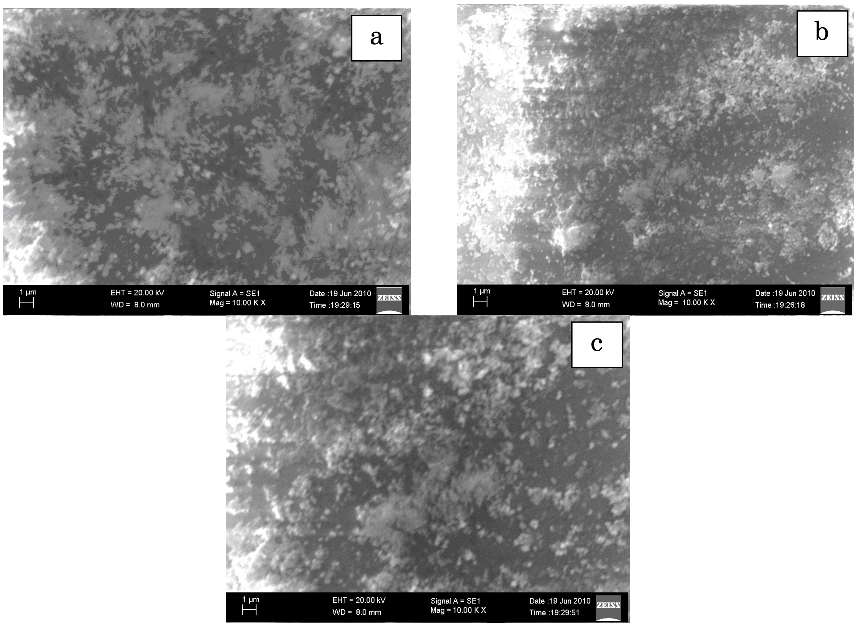


Fig. 4 *– SEM images of ZIS films prepared at pH = 9 (a) pH = 10 (b) and pH = 11 (c)*

3.4 Optical studies

The optical transmittance studies of ZIS films were carried out in the wavelength range, 300 - 2500 nm and the corresponding transmission spectra is shown in Fig. 5. The transmittance of ZIS films increased from 70 % to 80 % above the fundamental absorption edge with increase of pH. It could be observed from Fig. 5 that ZIS films had high transparency in the visible region and showed a steep absorption edge at a wavelength of about 400 - 450 nm, which indicates better crystallinity in ZIS films. It is noteworthy that the absorption edge shifted towards lower wavelength side with the increase of pH, indicating an increase of the energy band gap in ZIS films. The evaluated band gap varied from 2.15 eV to 2.64 eV with the change of pH in the range, 0 - 11, which is in good agreement with the behavior reported by Yadav et al. [9]. The larger energy band gap determined for the films grown at higher pH values might be attributed to the incorporation of hydroxyl groups in the films and/or due to the presence of secondary phases of Zn and Se in small quantity that were not observed in the XRD pattern.

to Contents, page 1

The evaluated refractive index varied in the range 1.85 - 1.73 with change of pH from 9 to 11 and the refractive index values obtained in the present were comparable with the values reported by Gordillo et al. [17] for thermal evaporated ZIS films.

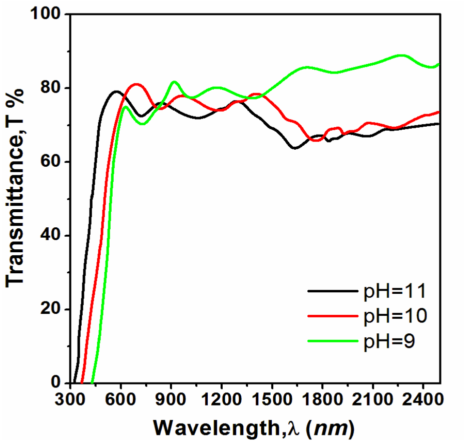


Fig. 5 *– Optical transmittance verses wavelength spectra of ZIS films*

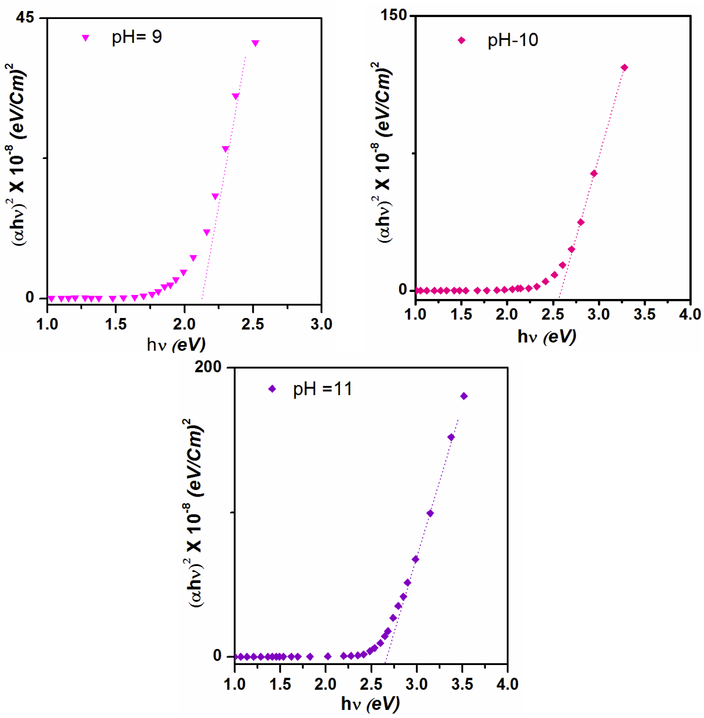


Fig. 6 *– Plots of $(\alpha h\nu)^2$ verses hν for ZIS films at different pH value*

to Contents, page 1

4. CONCLUSIONS

$ZnIn_2Se_4$ thin films have been deposited by simple and economic process, chemical bath deposition at different pH values. The pH of the deposition bath varied from 9 to 11 and the layers were grown at a fixed bath temperature (T_b) of 90 °C and a reaction time of 60 min using the ammonia and hydrazine hydrate as complexing agents. The XRD analysis revealed that the deposited films were polycrystalline, exhibiting tetragonal crystal structure and showed the (112) peak as the preferred orientation. SEM micrographs revealed that the samples had large number of granule like particles in different sizes. The optical transmittance of these samples was found to be > 75% in the visible region and the evaluated energy band gap varied from 2.15 eV to 2.64 eV with the change of pH value from 9 to 11.

REFERENCES

1. D. Lincot, R. Ortega-Borges, J. Vedel, M. Ruckh, J. Kesler, K.O. Velthaus, D. Harriskos, H.W. Schock, *Proc. 11th EC Photovoltaic Solar Energy Conf.*, 870 (1992).
2. T. Torndahl, C. Platzer-Bjorkman, J. Kessler, M. Edoff, *Prog. Photovolt. Res. Appl.* **15**, 225 (2007).
3. R.N. Bhattacharya, K. Ramanathan, L. Gedvilas, B. Keyes, *J. Phys. Chem. Solids* **66**, 1862 (2005).
4. A.E. Delahoy, J. Bruns, A. Ruppert, M. Akhtar, L. Chen, Z.J. Kiss, *Energy Photovoltaics*, Inc.Princeton, New Jersey, August, Nrel/sr-520-28786 (2000).
5. J. Luengo, N.V. Joshi, *Mater. Lett.* **26**, 47 (1996).
6. J. Filipowicz, N. Romeo, L. Trricone, *Solid State Commun.* **38**, 619 (1981).
7. H. Hahn, G. Frank, W. Klingler, A. StoИrger, G. StoИrger, Z. Anorg, *Allg. Chem.* **279**, 241 (1955).
8. M. Konagai, Y. Ohtake, T. Okamoto, *MRS Symp. Proc.* **426**, 153 (1996).
9. S.P. Yadav, P.S. Shinde, K.Y. Rajpure, C.H. Bhosale, *J. Phys. Chem. Solids* **69**, 1747 (2008).
10. L.I. Soliman, T.A. Hendia, *Radiat. Phys. Chem.* **50**, 175 (1997).
11. X. Sun, Y. He, J. Feng, J. Cryst., *J. Cryst. Growth* **312**, 48 (2009).
12. P. Babu, N. Revathi, P. Prathap and K.T. Ramakrishna Reddy, *Thin Solid Films*, (Accepted).
13. H.M. Zeyada, M.S. Aziz, A.S. Behairy, *Physica B* **404**, 3957 (2009).
14. B.D. Cullity, S.R. Stock, *Elements of X-ray Diffraction*, (Prentice Hall, Pearson: 2001).
15. C. Barret, T.B. Massalaski, *Structure of Metals*, 204 (Pergmon Oxford: 1980).
16. H.R. Moutinho, F.S. Hasoon, F. Abulfotuh, L.L. Kazmerski, *J. Vac. Sci. Technol. A* **13**, 2877 (1995).
17. G. Grodillo, C. Caldero, F. Rojas, *Rev. Mex. Fis.* **49**, 329 (2003).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.92-100

PACS numbers: 78.67.Bf, 75.75. + a, 61.72.uf

SYNTHESIS AND CHARACTERIZATION OF UNDOPED AND Co-DOPED SnO_2 NANOPARTICLES

R.N. Mariammal*, N. Rajamanickam, K. Ramachandran

School of Physics, Madurai Kamaraj University
Madurai – 21, India
* E-mail: marigayathrirn@yahoo.co.in

Undoped and Co-doped (1 and 3 at. %) SnO_2 nanoparticles were synthesized by a simple co-precipitation method. X-Ray diffraction data revealed that both undoped and doped samples crystallize in the tetragonal rutile phase with CoO phase in doped samples. The lattice parameters of doped samples calculated from XRD data do not vary much when compared to undoped one indicating that Co has not substituted the host lattice. The surface morphology investigated by SEM indicates the cluster formation in Co-doped (1 and 3 at. %) nanoparticles and the chemical composition of the samples were analyzed by energy dispersive spectroscopy (EDS). UV-Vis spectrum of undoped system showed absorption at 408 nm (3.04 eV), which is red shifted by 0.56 eV compared to bulk SnO_2 (3.6 eV) due to the cluster nature of the sample. The UV-Vis spectra of doped samples showed absorption in the visible region due to the formation of CoO phase, which is also evident from the XRD spectra. PL spectra showed characteristic UV emission at 409 nm and blue emission at 480 nm. The characteristic vibrational modes of SnO_2 were studied from FTIR analysis. EPR measurement of Co-doped (3 at .%) SnO_2 nanoparticles showed the paramagnetic behavior which may be attributed to the occupation of Co^{2+} ions in the interstitial site rather than the substitutional site. The absence of ferromagnetism is due to the high doping concentration of Co (> 1 at. %) and also due to the high annealing temperature which destroys the hyperfine splitting.

Keywords: *NANOPARTICLES, SEMICONDUCTORS, SCANNING ELECTRON MICROSCOPY, PHOTOLUMINESCENCE, MAGNETIC PROPERTIES.*

(Received 04 February 2011, in final form 18 March 2011)

1. INTRODUCTION

In recent years, diluted magnetic semiconductors (DMS) such as ZnO, TiO_2, and SnO_2 doped with transition metals (Co, Ni, V, Mn, Cr, and Fe) have attracted considerable interest due to their possibility of room temperature ferromagnetism (RTFM). Among these oxides, SnO_2 based DMS present special properties, such as giant magnetic moment and large coercivity [1]. SnO_2 is an n-type semiconducting oxide with a wide band gap of 3.6 eV. They have a wide range of applications in gas sensors [2, 3], optoelectronic devices [4], dye-sensitized solar cells [5], lithium batteries [6], and catalysts [7]. Many works reported RTFM in Co-doped SnO_2 nanoparticles. Still there are several controversies since some of the reports revealed the signature of paramagnetism and antiferromagnetism in these materials [8]. The oxygen stoichiometry plays an important role in RTFM in this system as reported by Punnose et al. [9]. Srinivas et al. [10] observed ferromagnetism in Co-doped (5 at.%) SnO_2 nanoparticles and found that the ferromagnetism increases

to Contents, page 1

with increasing annealing temperature. They argued that the ferromagnetic properties are also attributed to the nanometric size of the materials in addition to the surface diffusion of Co ions and the distribution of defects such as oxygen vacancies or vacancy clusters. Punnose et al. [11] have discussed that uniform dopant distribution and annealing temperature play an important role in Co-doped SnO_2 nanoparticles for RTFM. Few reports showed that ferromagnetism disappears in Co-doped SnO_2 nanoparticles beyond a particular dopant concentration and annealing temperature [12]. In addition, the segregation of Co atoms on the surface of SnO_2 was also expected to destroy the RTFM [11]. Liu et al. [1] studied magnetism in Co-doped SnO_2 nanocrystals which showed antiferromagnetic behavior. Since many reports showed different results, here undoped and Co-doped (1 and 3 at. %) SnO_2 nanoparticles are synthesized, structurally and optically characterized in order to study the nature of magnetism in these samples.

2. EXPERIMENT

2.1 Synthesis

A simple co-precipitation method is adopted to synthesize undoped and Co-doped (1 and 3 at. %) SnO_2 nanoparticles. In a typical synthesis, tin (IV) chloride ($SnCl_4$) (0.1 M) was dissolved in 250 ml of de-ionized water and was magnetically stirred with drop wise addition of ammonia (NH_3) solution to maintain the pH of the solution at 7. Finally, the resultant solution was stirred constantly for 30 minutes. The obtained precipitate was centrifuged, washed several times with de-ionized water in order to remove the chloride and other unreacted ions, and then dried in atmosphere at 60 °C to get SnO_2 nanoparticles. Co doping (1 and 3 at. %) is achieved by adding appropriate amount of cobalt chloride ($CoCl_2.6H_2O$) to the SnO_2 stock solution. Finally, the nanoparticles obtained were annealed at 600 °C for 4 h in atmosphere to get undoped and Co-doped (1 and 3 at.%) SnO_2 nanoparticles. Flowchart for the synthesis of these nanoparticles is given in Fig. 1.

2.2 Characterization

The structural analysis of the synthesized nanoparticles was performed by recording the X-Ray diffraction (XRD) spectrum at room temperature using X-Ray diffractometer (PANalytical X'Pert Pro) with Cu-Kα as the radiation source (wavelength: 1.54056 Å) at a step size of 0.02° over the 2θ range of 10° to 90°. The morphology of the samples was examined by scanning electron microscopy (Hitachi S-3400N, Japan) operating at an accelerating voltage of 20 kV. The elemental analysis was carried out by energy dispersive spectroscopy (EDS) (Nortan System Six, Thermo electron corporation Instrument Super DRY II, USA). The UV-Vis absorption spectrum was recorded at room temperature using UV-Vis absorption spectrometer (Shimadzu). Fourier transform infrared (FT-IR) spectra were measured using the KBr method on a Fourier transform infrared spectrometer (Shiraz) at room temperature in the range of 4000 - 400 cm^{-1} with a resolution of 1 cm^{-1}. Photoluminescence (PL) studies were carried out using a photoluminescence spectrophotometer (Varian Cary Eclipse) and the spectra were recorded in the range of 375 - 680 nm using xenon flash lamp as the excitation source with an excitation wavelength of 350 nm.

to Contents, page 1

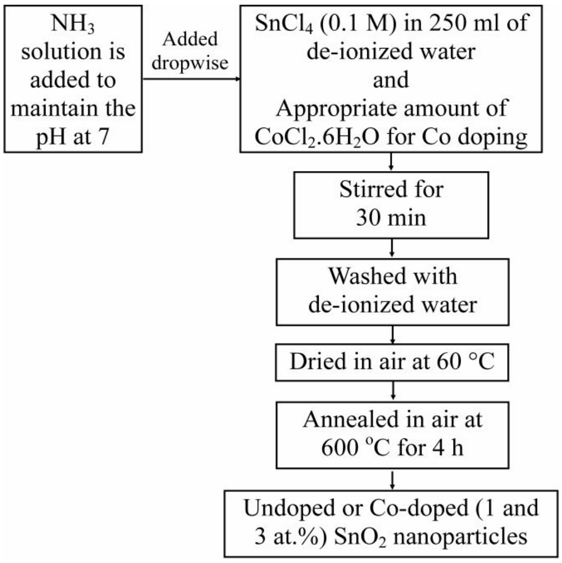


Fig. 1 *– Flowchart for the synthesis of undoped and Co-doped SnO_2 nanoparticles*

3. RESULTS AND DISCUSSION

3.1 X-Ray diffraction analysis

Fig. 2 shows the XRD spectra of undoped and Co-doped (1 and 3 at. %) SnO_2 nanoparticles. The spectra revealed that all the samples possess tetragonal rutile structure (JCPDS No. 41-1445). Eventhough Co-doped samples retain the rutile structure, they also show the presence of secondary phases of CoO.

Table 1 gives the lattice parameters of undoped and Co-doped SnO_2 nanoparticles. The lattice parameters of undoped SnO_2 nanoparticles are calculated from the XRD spectrum which are comparable with the JCPDS data and that of Co-doped samples do not show much variation compared to that of undoped one indicating that Co has not substituted the host lattice thus forming secondary phase of CoO, which is also evident from the appearance of extra peak corresponding to the hkl plane (002) of CoO in the XRD spectra of Co-doped samples. Also Co doping results in the broadening of the XRD peaks in the doped samples, which is due to the decrease in the crystallinity. Further by looking into the full width at half maximum of Bragg peaks, the average crystallite size of undoped and Co-doped SnO_2 nanoparticles is calculated from Scherrer's equation and is given in Table 1. The decrease in the crystallite size of doped samples when compared to undoped one is due to the decrease in the crystallinity of the doped samples.

Table 1 *– Results of XRD analysis on undoped and Co-doped (1 and 3 at. %) SnO_2 nanoparticles*

Samples	Crystallite size	a (Å)	b (Å)
SnO_2	12	4.70	3.18
SnO_2:Co (1 at.%)	9	4.71	3.18
SnO_2:Co (3 at.%)	8	4.69	3.18

to Contents, page 1

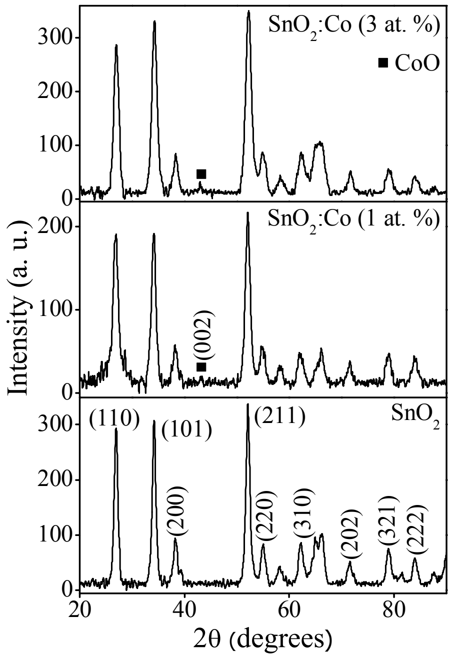


***Fig. 2** – XRD spectra of undoped and Co-doped (1 and 3 at. %) SnO_2 nanoparticles*

3.2 SEM analysis

The surface morphology of Co-doped (1 and 3 at.%) SnO_2 nanoparticles were studied by SEM analysis and the images are given in Fig. 3. The SEM images of both the samples show the formation of particles that agglomerated to form clusters. The cluster formation is due to annealing of the samples at 600 °C. Since SEM analysis is of low magnification, the individual particles are not clearly seen here. If we go for TEM analysis with high magnification, these particles can be clearly seen.

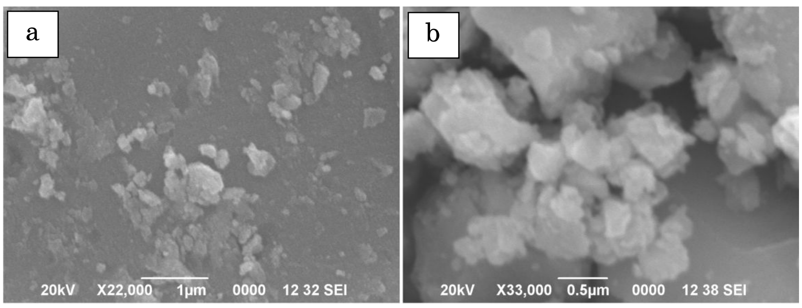


***Fig. 3** – SEM images of 1 at. % (a) and 3 at. % (b) Co-doped SnO_2 nanoparticles*

3.3 EDS analysis

The compositional analysis is necessary to monitor the concentration of the elements present in the sample. Fig. 4 (a and b) shows the typical EDS

to Contents, page 1

spectra of Co-doped (1 and 3 at. %) SnO_2 nanoparticles which revealed signals from zinc, oxygen, and cobalt elements only with no other impurities. But the peak corresponding to Pt is obtained in the spectra since Pt is sprayed on the sample while taking SEM for conductivity. Table 2 shows the compositions of Co-doped (1 and 3 at. %) SnO_2 nanoparticles estimated from EDS analysis. Whenever a material is doped, it is necessary to see whether the dopant has entered the host system completely or partially. From the EDS analysis, it has been observed that the Co concentration in the doped samples is less than that of the starting solution. Even though we intended to dope 1 and 3 at. % of Co, only 0.31 and 0.65 at. % has actually entered the host system and the remaining less energetic molecules would be washed away during synthesis.

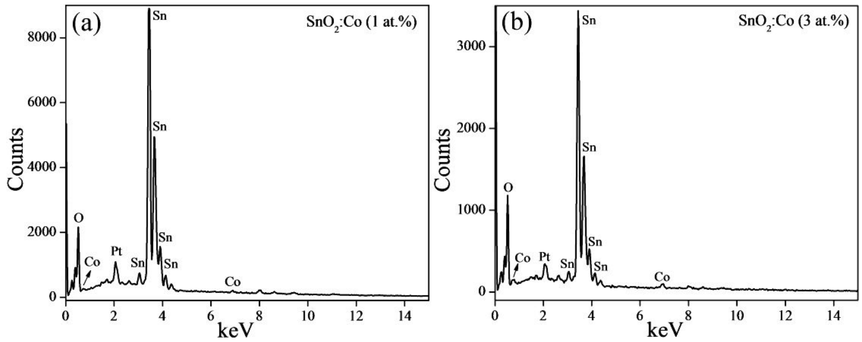


***Fig. 4** – EDS spectra of 1 at. % (a) and 3 at. % (b) Co-doped SnO_2 nanoparticles*

***Table 2** – The chemical compositions of Co-doped (1 and 3 at. %) SnO_2 nanoparticles*

Samples	Sn (at. %)	O (at. %)	Co (at. %)
SnO_2:Co (1 at.%)	21.45	78.23	0.31
SnO_2:Co (3 at.%)	15.99	83.36	0.65

3.4 UV-Vis analysis

The optical absorption spectra of undoped and Co-doped (1 and 3 at. %) SnO_2 nanoparticles are given in Fig. 5 and the spectrum of undoped sample shows absorption at 408 nm (3.04 eV) which is red shifted from the bulk (3.6 eV) due to the cluster nature of the sample. Co-doped samples do not exhibit a sharp absorption edge probably due to Co d states extending into the band gap region resulting from the overlapping of orbitals [10]. The absorption spectra of Co-doped samples also show visible absorption, which is due to the presence of CoO phase in these samples thus confirming that Co has not entered the host system.

3.5 Photoluminescence measurements

Fig. 6 shows the PL spectra of undoped and Co-doped (1 and 3 at. %) SnO_2 nanoparticles. PL measurements were carried out to understand the emission properties and the surface defects. Room temperature PL emission spectra of all the samples were recorded at an excitation wavelength of 350 nm which is in agreement with the reported value. The spectra show characteristic

to Contents, page 1

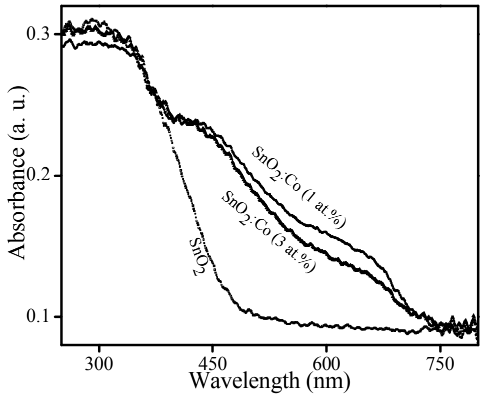


***Fig. 5** – UV-Vis absorption spectra of undoped and Co-doped (1 and 3 at.%) SnO_2 nanoparticles*

UV emission at 409 nm and this emission peak is due to all the luminescent centers, such as nanocrystals and intrinsic defects [13]. With the introduction of Co, the decrease in UV emission was attributed to the increased intrinsic defects of nanoparticles [13]. It is interesting to note that the emission spectral intensity of Co-doped (3 at. %) SnO_2 nanoparticles increases when compared to Co-doped (1 at. %) SnO_2 nanoparticles. This increased intensity is due to the decrease in the particle size as shown in Table 1 and this behavior may be attributed to an increase in the number of luminescence centers by increasing the ratio of surface area [10]. Apart from the UV emission, the spectra also show broad blue emissions at 480 and 492 nm that might be due to the electron transition mediated by defect levels in the band gap [13]. With the increase of Co concentration, the emission peaks broaden and this is mostly due to the decrease and non-uniformity of the particle size, which resembles with UV-Vis absorption studies.

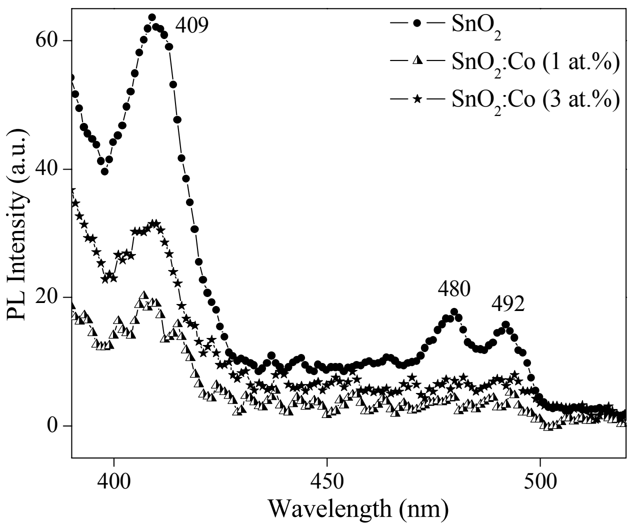


***Fig. 6** – PL spectra of undoped and Co-doped (1 and 3 at. %) SnO_2 nanoparticles*

to Contents, page 1

3.6 FTIR analysis

Fig. 7 shows the FTIR spectra of the synthesized samples. The transmission band at 514 cm $^{-1}$ is assigned to the Sn-O terminal bond of SnO_2 and that at 615 cm $^{-1}$ is assigned to the O-Sn-O bridging bond [14]. The principal peak at 1629 cm $^{-1}$ corresponds to the strong asymmetric stretching of C=O bond. The band observed at 2359 cm $^{-1}$ is assigned to the existence of CO_2 molecule in air. Vibrational mode observed at 2923 cm $^{-1}$ is due to C-H stretching vibration. The characteristic band at 3430 cm $^{-1}$ corresponds to the stretching vibration of O–H groups.

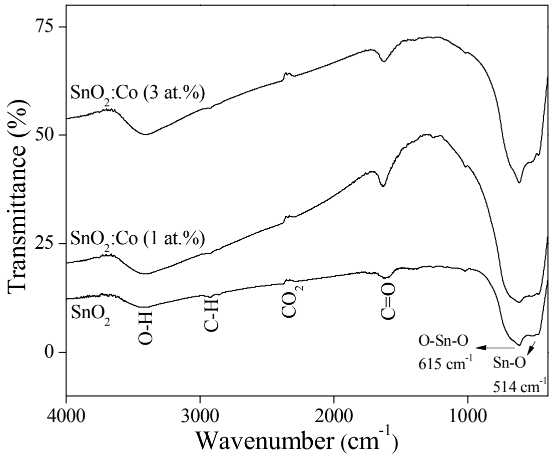


***Fig. 7** – FTIR spectra of undoped and Co-doped (1 and 3 at. %) SnO_2 nanoparticles*

3.7 EPR analysis

Fig. 8 shows the EPR spectrum of Co-doped (3 at. %) SnO_2 nanoparticles measured at room temperature. The spectrum resembles with the result obtained by Misra et al. [12]. They showed that this behavior is due to the incorporation of Co^{2+} ions with effective spin $S = 1/2$ in the interstitial site of SnO_2 nanoparticles. The expected hyperfine splitting terms corresponding to the ferromagnetism are not obtained here thus confirming the absence of ferromagnetism in this sample but exhibits a paramagnetic behavior [12]. Misra et al. [12] observed that the ferromagnetic behavior is lost when the doping concentration is greater than 1 at. % and also for samples annealed at temperature greater than 350 °C. Here the samples were annealed at 600 °C which may introduce randomly distributed defects as confirmed by PL emission spectra, which enhances disorder of the crystalline field at Co^{2+} sites, thereby increasing the hyperfine linewidth. Increasing the linewidth suppresses hyperfine splitting which further destroys the ferromagnetism. Also the SEM results showed the segregation of nanoparticles on the surface which is also responsible for the destruction of RTFM [11].

to Contents, page 1

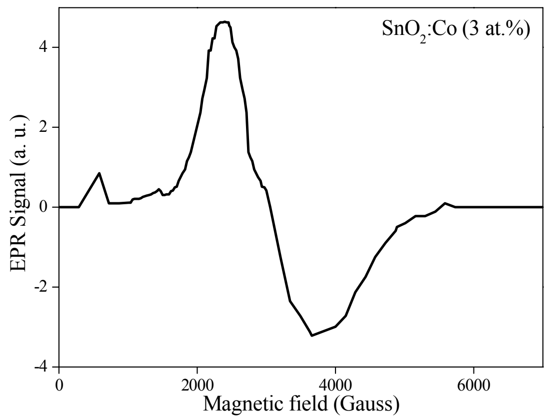


***Fig. 8** – EPR spectrum of Co-doped (3 at. %) SnO_2 nanoparticles*

Lande-g factor is calculated using

$$g = \frac{h\nu}{\mu_B H}, \quad (1)$$

where h is Planck constant, ν is the frequency, μ_B is the Bohr magneton, and H is the applied magnetic field in G. The value of g is found to be 2.819 for SnO_2:Co (3 at. %) nanoparticles.

4. CONCLUSION

Undoped and Co-doped (1 and 3 at.%) SnO_2 nanoparticles were synthesized by chemical co-precipitation method. The XRD analysis showed that the synthesized SnO_2 nanoparticles have tetragonal rutile structure and revealed the formation of secondary phase of CoO due to Co doping. The lattice parameters of doped samples do not vary much compared to undoped one. The appearance of CoO phase in the XRD spectra and the invariance of lattice parameters of doped samples confirmed that Co has not entered the substitutional site of the host system, which is further confimed by UV-Vis studies. The surface morphology and chemical composition were examined using SEM with EDS analysis. The absorption spectra of undoped sample showed absorption at 408 nm, which is red shifted from bulk and that of doped samples showed absorption in the visible region (due to CoO phase) without any sharp absorption peak. PL studies showed strong UV emission at 409 nm whose intensity decreased with the addition of Co which is due to the introduction of more intrinsic defects. Further, the vibrational modes were studied from FTIR analysis. EPR spectrum showed that Co-doped (3 at. %) SnO_2 nanoparticles exhibit paramagnetic behavior at room temperature.

ACKNOWLEDGEMENTS

The authors acknowledge DST Project on Nanotechnology and UGC-DRS & UPE for the financial support.

to Contents, page 1

REFERENCES

1. C.M. Liu, X.T. Zu and W.L. Zhou, *J. Phys.: Condens. Matter* **18,** 6001 (2006).
2. Y. Wang, X. Jiang and Y. Xia, J. Am. *Chem. Soc.* **125,** 16176 (2003).
3. A.A. Firooz, A.R. Mahjoub and A.A. Khodadadi, *Sensor. Actuator. B* **141,** 89 (2009).
4. F. Gu, S.F. Wang, M.K. Lu, Y.X. Qi, G.J. Zhou, D. Xu and D.R. Yuan, *Opt. Mater.* **25,** 59 (2004).
5. S. Ferrere, A. Zaban and B.A. Gsegg, *J. Phys. Chem. B* **101,** 4490 (1997).
6. V. Subramanian, K.I. Gnanasekar and B. Rambabu, *Solid State Ionics* **175,** 181 (2004).
7. S.R. Stampfl, Y. Chen, J.A. Dumesis, Ch. Niu, C.G. Hill, *J. Catal.* **105,** 445 (1987).
8. A. Bouaine, N. Brihi, G. Schmerber, C.U. Bouillet, S. Colis and A. Dinia, *J. Phys. Chem. C* **111,** 2924 (2007).
9. A. Punnoose, J. Hays, V. Gopal and V. Shutthanandan, *Appl. Phys. Lett.* **85,** 1559 (2004).
10. K. Srinivas, M. Vithal, B. Sreedhar, M.M. Raja and P.V. Reddy, *J. Phys. Chem. C* **113,** 3543 (2009).
11. A. Punnoose, M.H. Engelhard and J. Hays, *Solid State Commun.* **139,** 434 (2006).
12. S.K. Misra, S.I. Andronenko, K.M. Reddy, J. Hays and A. Punnoose, *J. Appl. Phys.* **99,** 08M106 (2006).
13. M. Fang, X.T. Zu, Z.J. Li, S. Zhu, C.M. Liu, L.M. Wang and F. Gao, *J. Mater. Sci.: Mater. Electron* **19,** 868 (2008).
14. O.N. Gavrilenko, E.V. Pashkova and A.G. Belous, *Russ. J. Inorg. Chem.* **52,** 1835 (2007).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2010) No1, P.101-105

PACS numbers: 61.05.Cp, 61.72.Dd, 61.72.Ff,07.85.Jy, 61.72.Hh, 61.72.Nn

SUBSTRATE TEMPERATURE EFFECT ON STRUCTURAL PROPERTIES OF Bi_2Te_3 THIN FILMS

B.S. Jariwala, D.V. Shah, Vipul Kheraj

Department of Applied Physics,
S.V. National Institute of Technology,
Ichchha Nath, Piplod, 395 007, Surat, India
E-mail: j.bhakti@ashd.svnit.ac.in

Structural properties of Bi_2Te_3 thin films, thermally evaporated on well-cleaned glass substrate at different substrate temperature, are reported here. X-ray diffraction was carried out for the structural characterization. XRD pattern of the films exhibits preferential orientation along the [0 1 5] *direction for the films of all the substrate temperature together with other supported planes* [2 0 5] *&* [1 1 0]*. All deposition conditions like thickness, deposition rate and pressure were maintained throughout the experiment. X-ray diffraction lines confirm that, the grown films are polycrystalline in nature with the hexagonal crystal structure. The effect of substrate temperature on these parameters have been investigated and reported in this paper. Various structural parameters such as lattice constants, grain size, micro strain, number of crystallites, stacking fault and dislocation density were calculated using X-ray diffraction analysis*

Keywords: *BISMUTH TELLURIDE THIN FILMS, X-RAY DIFFRACTION, STRUCTURAL PARAMETERS, DISLOCATION DENSITY, MICRO STRAIN, STACKING FAULT.*

(Received 04 February 2011, in final form 20 March 2011)

1. INTRODUCTION

The V_2-VI_3 (V = Bi, VI = Se, Te) binary compounds and their pseudo binary solid solutions are highly anisotropic and crystallize into homologous layered structure parallel to c-axis and are known to find applications ranging from photoconductive targets in TV cameras to IR spectroscopy [1, 2]. Among these Bi_2Te_3 is the most potential candidates for thermoelectric devices such as thermoelectric generators, thermocouples, thermo coolers and IR sensors with the best figure of merit near room temperature [3-6]. It also has applications in electronics, microelectronic, optoelectronic and electro-mechanical devices. It is a p-type semiconductor with a direct band gap of 0.21 eV and melting point 585 °C. The high ratio of the electrical conduction to the thermal conductivity makes bismuth telluride a good thermoelectric material [7-10]. There have been various studies on the optical and electrical properties of Bi_2Te_3 thin films. The dependency of thickness of film on structural properties were studied by Sathyamoorthy [11]. Here, the present authors report the structural variation in properties with the substrate temperature for a film thickness of 1000 Å.

to Contents, page 1

2. EXPERIMENTAL DETAILS

Bi_2Te_3 compound was synthesized from its constituent elements of 99.999% purity. Bi and Te were sealed in quartz ampoule at a pressure of 10^{-5} torr in stoichiometric proportion. The ampoule was subjected to the alloy mixing furnace at the temperature of 620 °C. The temperature of the furnace was raised at the rate of 50 °C/hr and maintained for 24 hours. During this, the ampoule was continuously rocked and rotated for uniform mixing and homogeneity of the melt. Bi_2Te_3 thin films were deposited on the well-cleaned glass substrates held at different substrate temperature by thermal evaporation technique under the base vacuum of 10^{-6} torr. Here, the Bi_2Te_3 powder was taken from the ingot, in order to achieve better composition condition. The rate of deposition and the thickness of the film were measured to be 1 Å/s and 1000 Å respectively using the quartz crystal monitor.

The chemical composition of the ingot was obtained by EDAX(Philips ESEM). The X-ray diffraction patterns of the deposited Bi_2Te_3 thin films were recorded in the range of 20 - 70 degree with the help of Philips X-ray diffractometer using CuK_α radiation.

3. RESULT AND DISCUSSION

3.1 Compositional characterization

The typical EDAX, carried out for the synthesized ingot is shown in the Figure 1. The analysis confirms the stoichiometry of the synthesized compound with Bi and Te in the atomic percentage.

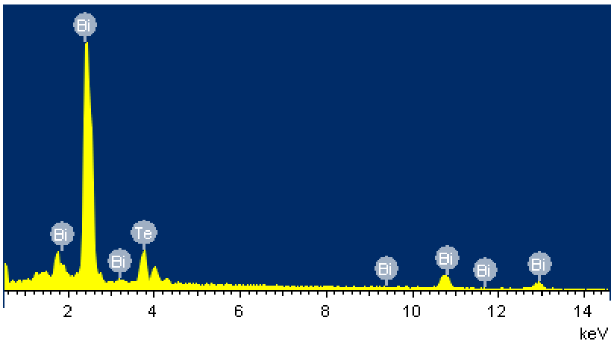


Fig. 1 *– EDAX of synthesized Bi_2Te_3 ingot*

3.2 Structural characterization

X-ray diffraction was carried out on evaporated Bi_2Te_3 thin films evaporated at different substrate temperatures and are presented in Figure 2. The thickness of the films was 1000 Å. They exhibited peaks at $2\theta = 22.54°$, 27.64°, 37.89°, 40.14° which corresponded to diffractions of (2 0 5), (0 1 5), (0 1 8) and (0 1 1) plane of hexagonal phases, respectively, indicating the polycrystalline nature of the films. The plane indices are obtained by comparing with the JCPDS # 020524 and #150863. The (0 1 5) hexagonal plane reflects Bi_2Te_3 phase present in the grown thin film.

to Contents, page 1

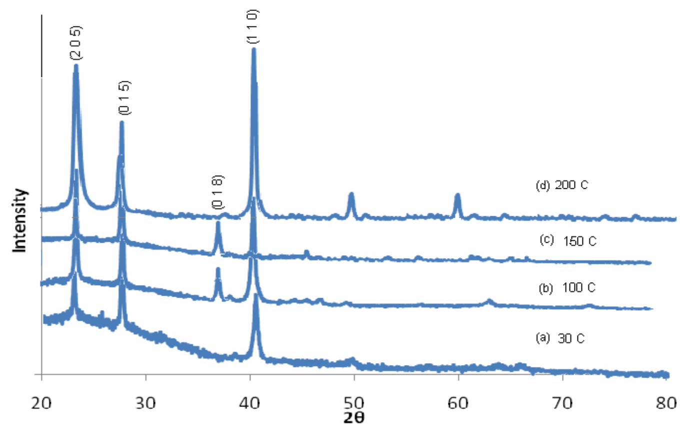


Fig. 2 *– XRD of Bi_2Te_3 thin film of substrate temperature (a)* 30 °C *(b)* 100 °C *(c)* 150 °C *(d)* 200 °C

The lattice parameter 'a' and 'c' are determined by using the relation for hexagonal crystal structure for the diffraction plane (1 1 0) and found to be a = 4.384 Å and c = 30.424 Å which are close to the reported values [11, 12]. The various structural parameters for Bi_2Te_3 thin films are calculated using the relevant formula [11, 13, 14] and are represented in Table 1.

Microstrain (ε):

$$\varepsilon = \frac{\beta_{2\theta}\cos\theta}{4}$$

Estimation of number of crystallites:

$$N = \frac{t}{D^3}$$

Dislocation density:

$$\rho = \frac{15\varepsilon}{aD}$$

Stacking fault:

$$3\alpha + 3\beta = \frac{\beta_{2\theta}\pi^2c^2}{360ld^2\tan\theta} \quad \text{(for even } l\text{)}$$

$$3\alpha + \beta = \frac{\beta_{2\theta}\pi^2c^2}{360ld^2\tan\theta} \quad \text{(for odd } l\text{)}$$

Where, $\beta_{2\theta}$ is the full width at half maximum, λ is the wavelength of CuKα radiation (1.54 Å), D is the mean crystallite size, t is the thickness of the film, ε is the microstrain, α is the deformation fault probability and β is the growth fault probability.

The grain size, micro strain, dislocation density and No. of crystallite of the deposited film at various substrate temperature was shown in Figure 3a, 3b, 3c, 3d respectively. As shown in Figure 3a, the grain size rapidly increases with the temperature upto 100 °C. No major change have been seen between 100 °C to 150 °C where as it again decreases with the increasing temperature. The larger grain size within this temperature range indicates the formation of good quality films.

to Contents, page 1

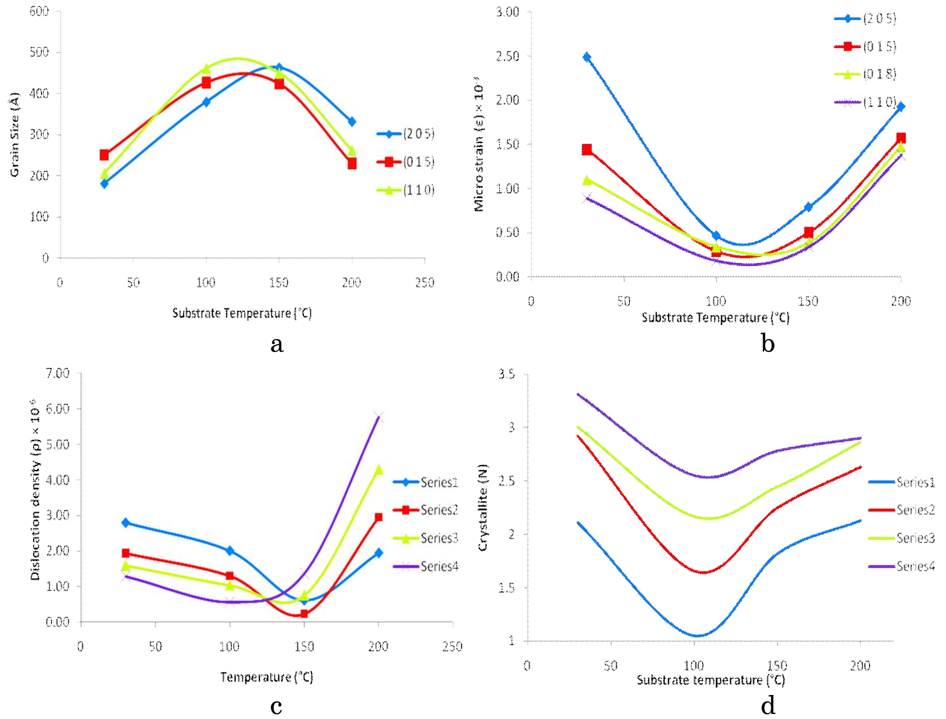


Fig. 3 *– Variation of (a) grain size (b) micro strain (c) dislocation density (d) No. of crystallite with substrate temperature*

Table 1 *– Structural parameters of thermally evaporated Bi_2Te_3 thin film at various substrate temperatures*

Subs. Temp. (°C)	$h\ k\ l$	D (Å)	$\varepsilon 10^{-3}$	$N \cdot 10^{16}$ (cm^{-3})	$\rho 10^{-6}$(cm^{-2})	α	β
30	2 0 5	181.00	2.49	2.11	8.40	0.10812	– 0.10097
	0 1 5	251.70	1.44	2.92	0.194		
	0 1 8	236.22	1.10	3.01	0.160		
	1 1 0	206.57	0.89	3.31	0.130		
100	2 0 5	380.20	0.464	1.05	2.01	0.09793	– 0.08078
	0 1 5	426.79	0.285	1.66	7.60		
	0 1 8	448.36	0.338	2.17	0.124		
	1 1 0	462.01	0.183	2.55	5.73		
150	2 0 5	464.21	0.792	1.82	6.15	0.06782	– 0.04400
	0 1 5	424.40	0.500	2.25	2.33		
	0 1 8	438.67	0.385	2.45	7.57		
	1 1 0	450.84	0.340	2.78	0.137		
200	2 0 5	332.19	0.953	2.13	0.195	0.15142	– 0.12761
	0 1 5	230.48	1.57	2.63	0.295		
	0 1 8	245.25	1.48	2.87	0.430		
	1 1 0	262.52	1.38	2.90	0.577		

to Contents, page 1

It is observed from Figure 3b, c and d that the micro strain, dislocation density and number of crystallite decrease up to 100 °C and then become steady up to 150 °C. These parameters again increase as the substrate temperature is raised further. The initialization of crystalline grains with considerable size was observed at 100 °C substrate temperature, as evident from Figure 2 and 3a. The maximum grain size was obtained for the substrate temperature between 110 °C and 150 °C. The structural properties depend upon arrangement of the evaporated atoms or molecules arriving on the substrate surface which possess larger kinetic energy at higher substrate temperatures. The appropriate substrate temperature gives large surface mobility and provides optimum diffusion distance of the evaporated atoms.

4. CONCLUSION

Bi_2Te_3 compound was synthesized from its constituent elements in a stoichiometric proportion and the maximum grain size 464.21 Å was obtained at 150 °C substrate temperature. Minimum micro strain 0.285 as well as dislocation density 0.124 were obtained for (0 1 8) plane at 100 °C. Whereas, the deformation fault probability α was reported 0.06782, higher than the growth fault probability β, − 0.04400, which indicate the lowest stacking fault probability at the substrate temperature of 150 °C. Thus, 100 °C to 150 °C is the optimum range of substrate temperature for the growth of high quality Bi_2Te_3 thin films using evaporation technique.

REFERENCES

1. D. Arivuoli, F.D. Gnanam, P. Ramasamy, *J. Mater. Sci. Lett.* **7**, 711 (1988).
2. M. Stolzerm, M. Stordeur, H. Sobotta, V. Riede, *phys. status solidi B* **138**, 259 (1986).
3. D.M. Rowe, C.M. Bhandari, *Modern thermo electronics* (London: Holt Rinehbort & Winston: 1981).
4. C.H.L. Goodman, *Mater. Res. Bull.* **20**, 237 (1985).
5. L. Jansa, P. Lostak, J. Sramkova and J. Horak, *J. Mater. Sci.* **27**, 6062 (1992).
6. B. Roy, B.R. Chakraborty, R. Bhattacharya, A.K. Dutta, *Solid State Commun*. **25**, 937 (1978).
7. H.W. Jeon, H.P. Ha, D.B. Hyun and J.D. Shim, *J. Phys. Chem. Solids* **52**, 579 (1981).
8. T. Plechacek, J. Navrotil and J. Horak, *J. Solid State Chem.* **165**, 35 (2002).
9. S.R. Bhavsar, G.R. Pandya, P.H. Soni, C.F. Desai, N.R. Shah, *Optic & Opto-electronics* **2**, 1344 (1998).
10. N. Sakia, T. Kajiwara, K. Takemura, S. Minomura, Y. Fuji, *Solid State Commun.* **40**, 1045 (1981).
11. R. Sathyamoorthy, J. Deepa, *J. Phys. Chem. Solids.* **68**, 111 (2007).
12. M. Francombe, *Philos. Magn.* **10**, 989 (1994).
13. B.D. Cullity, *Elements of X-ray diffraction* (USA, Addison-Wesley: 1956).
14. J. Deepa, R. Sathyamoorthy, S. Velumani, A. Subbarayan, K. Natarajan P.J. Sebastian, *Sol. Energ. Mat. Sol. C.* **81**, 305 (2004).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.106-110

PACS numbers: 77.55. + f, 81.20.Fw, 33.20.Ea

DEPOSITION AND SURFACE MODIFICATION OF LOW-K THIN FILMS FOR ILD APPLICATION IN ULSI CIRCUITS

Y.S. Mhaisagar, B.N. Joshi, A.M. Mahajan

Department of Electronics, North Maharashtra University,
Jalgaon – 425001 (M.S.), India
Email: ammahajan@nmu.ac.in

The low-k thin films have been deposited successfully by sol gel technique using tetraethylorthosilicate (TEOS) precursor and the surface of deposited thin films have been modified by wet chemical treatment using trimethylcholorsilane (TMCS) and hexane solution with 15 % *volume ratio to remove the hydroxyl groups from the surface of deposited low-k thin films. The characterization of the as deposited and surface modified low-k thin films has been carried out by Ellipsometer, Fourier transform infrared (FTIR) spectrometer, and contact angle meter. For the determination of the dielectric constant of the deposited thin film the metal – insulator-semiconductor (MIS) structure was formed by depositing the Aluminium (Al) metal on the low-k thin film. Further the capacitance-voltage curve of the MIS structure has been obtained at* 1 *MHz frequency. The dielectric constant of the as deposited thin film is found to be* 2.15*. The lowering of O-H peaks and appearance of* CH_3 *peaks in FTIR spectra confirms the surface modification of* SiO_2 *films. The contact angle of the deposited thin film is changed from* 83.3° *to* 104° *after surface modification that validates the transformation of thin film surface from hydrophilic to hydrophobic after the surface modification treatment.*

Keywords: *LOW-K, TMCS, SURFACE MODIFICATION, HYDROPHOBIC, CV, CONTACT ANGLE.*

(Received 04 February 2011, in final form 19 March 2011)

1. INTRODUCTION

The introduction of the advanced fabrication processes causes the new revolution in the technology that helps in scaling the device size towards the nanoregime. The scaling in the device size towards the nanoregime results in high speed and reliable devices. But this miniaturization of devices in ultra large scale (ULSI) integrated circuits with conventional materials limits their speed due to the properties of materials like resistivity, mechanical strength etc. Thus, an Al metal in ULSI is replaced with Copper (Cu) for backend of line (BEOL) application due to the low resistivity of Cu interconnect which minimizes the resistance (R) in RC (C-capacitance) delay of interconnect. The reduction in interconnect capacitance has also got significant attention alongwith lowering in interconnect resistance in order to lower the RC delay. To lower the interconnect capacitance it is essential to minimize the dielectric constant of insulator between interconnects used as interlayer dielectric (ILD) in ULSI circuits. As per the recent ITRS the dielectric constant for the 32 nm technology should be less than 2.5. The dielectric constant can be reduced by lowering the density of the materials

to Contents, page 1

(by introducing the porosity) or by introducing lower polarized bonds. But by introducing the lower polarized bonds it is difficult to achieve the ultra low-k value of thin films. Thus, the another option to lower the density of the materials is to incorporate porosity in the film [1-3].

Two important techniques being used to deposit the low-k thin film are the PECVD and the sol gel out of which, the sol-gel is the best method because of its ability to introduce the high porosity with controlled pore size [4-5]. The sol-gel deposited low-k thin films contains the large number of hydrophilic groups (−OH bonds) on its surface that may detoriate the dielectric constant of the film. Such hydrophilic groups can not be fully eliminated during the annealing processes. Thus to prevent change in dielectric constant of the porous thin film the surface should be modified from hydrophilic to hydrophobic [6].

From the literature survey it is observed that wet chemical treatment, Plasma treatment and UV treatment are mostly used to enhance the properties of the porous low-k thin film [7-10].

In present work, the wet chemical treatment method has been used for the surface modification of low-k thin films. The surface modifications were carried out by using TMCS / Hexane as a modifying agent. During the surface modification processes hydroxyl groups present on surface of films gets replaced by trimethylsilyl (TMS) groups from the TMCS preventing silica condensation reactions during drying. These low-k thin film have been characterized by Ellipsometer, FTIR and contact angle measurement setup and Capacitance-voltage analyzer before and after surface modification process.

The second section of this paper explains the experimental part, in the third section the results are discussed and fourth section concludes the paper.

2. EXPERIMENTAL DETAILS

The precursor solution was prepared by mixing TEOS with ethanol, deionised water and acid catalyst at room temperature for deposition of porous SiO_2 low-k thin films with a molar ratio of 1:4:2:0.1 of TEOS: Ethanol: H_2O: HF respectively. The mixture solution was stirred for 1 hour at ambient temperature. The prepared sol was then dispensed on p-type Si $\langle 100 \rangle$ substrate before the gel point is reached and spun by spin coater. The spin coated, thin films were then heated at 100° C at room temperature. Then, the films were modified by chemical treatment method using the TMCS ($(CH)_3$ SiCl) as a sylinating agent with 15 % in hexane for 5 hours. The surface modified films were dried on hot plate at 80 °C for 30 minute and then finally annealed in a closed furnace at 300 °C for an hour. The films were characterized before and after the surface modification using by Ellipsometer (Philips SD 1000) for thickness measurement, FTIR (Nicolet 380) for chemical bonding analysis and contact angle meter (GBX make) for confirmation of hydrophobic surface. The dielectric constant of deposited film was determined by forming MIS structure using C-V analyzer.

3. RESULTS AND DISCUSSION

The as-deposited and surface modified SiO_2 thin films have been characterized by Ellipsometer having He-Ne laser of wavelength 632.8 nm. The average thickness of all the films is about 200 nm. The FTIR

to Contents, page 1

characterization of deposited and surface modified thin films have been carried out in the range of 400 - 4000 cm^{-1} with resolution of 4 cm^{-1} and scan rate of 128 to obtain information about surface bonding characteristics and evidence of surface modification. The FTIR absorption spectra of as deposited and surface modified films are presented in figure 1. The peaks at 441 and 805 cm^{-1} are identified as the rocking and bending vibration modes of Si-O-Si respectively and the peak at 1077.9 cm^{-1} corresponds to the Si-O-Si stretching vibration, which confirms the formation of Si-O-Si network in the films. The broad peak at 3416.9 cm^{-1} is due to O-H stretching vibration appeared in as deposited film just before the surface modification illustrates the hydrophilic nature of the film. After the surface modification of the deposited film using 15 % of TMCS solution in hexane the peak at 3416.9 cm^{-1} due to hydroxyl group gets diminished in FTIR spectra. The appearance of carbon peak in FTIR spectra of the surface modified film is due to replacement of -OH group by -CH group from TMCS [11]. The peak at 2960 and 2921 cm^{-1} corresponds to the asymmetric and symmetric stretching of CH_3 group respectively, while the peak at 2855 cm^{-1} represents the symmetric stretching of C-H in CH_2 group [12]. The intense peak at 1259 cm^{-1} is attributed to the Si-C bond [6]. This incorporation of carbon group into the modified film indicates the formation of hydrophobic nature of the film.

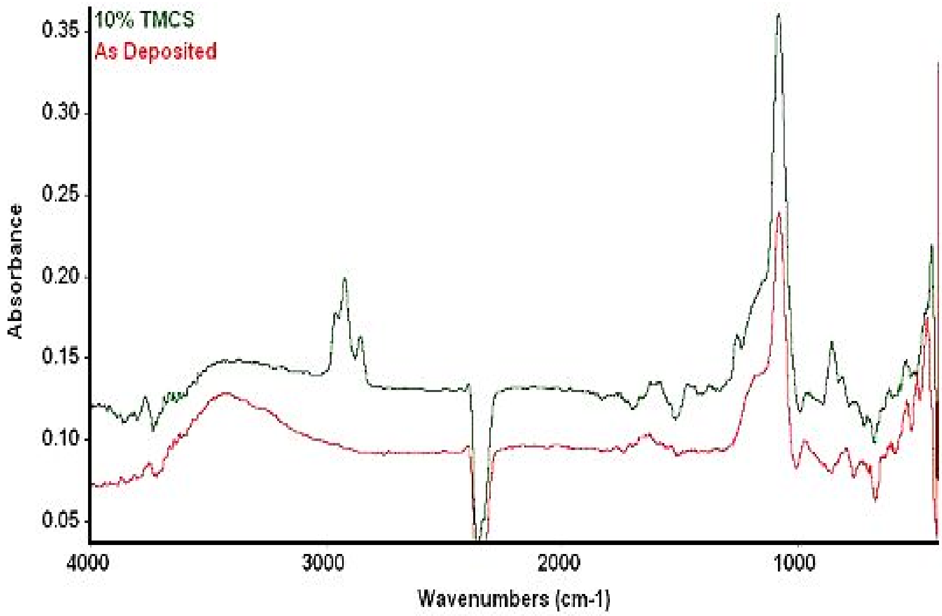


***Fig. 1** – FTIR spectra of as deposited and surface modified thin film*

The dielectric constant of the as deposited thin film has been carried out by forming MIS capacitor by depositing Al metal on deposited thin film as a gate electrode and on lower side of the Si substrate for formation of second contact. The top contact area of the metal film was 2.48×10^{-2} cm^2. The capacitance – voltage characteristics has been carried out at 1 MHz frequency at IIT Mumbai. The dielectric constant of the deposited film is determined

to Contents, page 1

from the equation 1 [13] using the accumulation capacitance of C-V curve as shown in Figure 2. The k value of the film before surface modification is determined to be 2.15.

$$k = \frac{C_{ap} \cdot t}{\varepsilon_o \cdot A}, \tag{1}$$

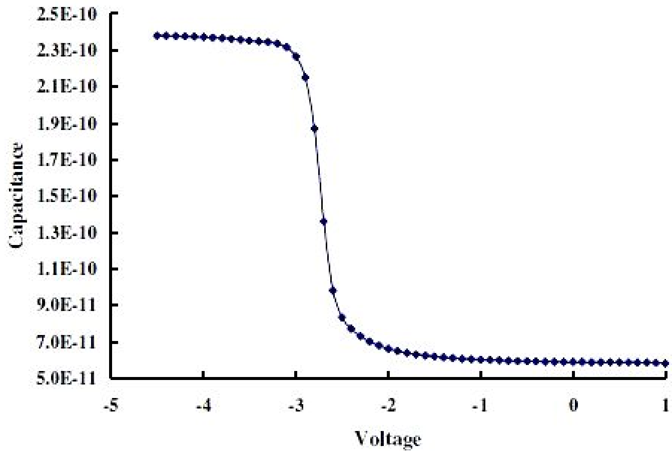


Fig. 2 – *Capacitance Voltage curve at* 1 *MHz frequency*

The hydrophobic nature of the surface modified low-k thin film has been determined by the contact angle measurement of the water droplet on the surface of the low-k thin film using contact angle meter. In this process, the drop of water is added on top surface of the film and the water droplet is photographed. The contact angle has been calculated directly from software by using young's equation. The contact angle of as deposited and surface modified film is shown in Figure 3a and b respectively. The contact angle of as deposited and surface modified film is observed to be 83.3° and 104° respectively. From the measured contact angle value it confirms that the hydrophilic surface of the film has became hydrophobic after wet chemical treatment, as the contact angle after modification is more than 90° [14].

a b

Fig. 3 – *Contact angle figure (a) hydrophilic surface (θ = 83.3°), (b) hydrophobic surface (θ = 104°)*

4. CONCLUSIONS

The SiO_2 low-k thin films of 200 nm thickness have been deposited successfully by sol gel spin coating technique using tetraethylorthosilicate (TEOS) as a source of Si. The hydrophilic nature of the film has been improved by silylation method. The silylation of the low-k thin film have been carried out by using TMCS/Hexane as surface modifying agent. The FTIR peaks of the surface modified low-k thin film appearing at 852 cm^{-1} and 2959 cm^{-1} confirms the replacement of the hydrophilic groups. The contact angle of 104° confirms the surface modification of the low-k thin film from hydrophilic to hydrophobic by TMCS silylation. Such hydrophobic thin films with low dielectric constant of 2.15 are suitable for ILD applications in ULSI circuits.

ACKNOWLEDGEMENTS

Authors would like to acknowledge with thanks the financial assistance provided by CSIR, New Delhi to carry out this work (vide letter No.03 (1127)/08/EMR-II) and authors are also greatful to the CEN, IITB (under INUP at IITB which have been sponsored by DIT, MCIT, and Government of India) for providing the characterization facility of contact angle and C-V measurements.

REFERENCES

1. K. Hamioud, V. Arnal, A. Farcy, V. Jousseaume, A. Zenasni, B. Icard, J. Pradelles, S. Manakli, Ph. Brun, G. Imbert, C. Jayet, M. Assous, S. Maitrejean, D. Galpin, C. Monget, *Microelectron. Eng.* **87,** 316 (2010).
2. D. Moore, R. Carter, H. Cui, P. Burke, P. McGrath, and S.Q. Gu, D. Gidley, H. Peng, *J. Vac. Sci. Technol. B* **23,** 332 (2005).
3. C. Dubois, A. Sylvestre, H. Chaabouni, A. Farcy, *Microelectron. Eng.* **87,** 333 (2010).
4. M.R. Baklanov, K. Maex, *Philos. T. Roy. Soc. A* **364,** 201 (2006).
5. Z.-W. He, D.-Y. Xu, X.-H. Jiang, Y.-Y. Wang, *Chinese Phys. B* **17,** 3021 (2008).
6. J.-T. Luo, W.-F. Wu, H.-C. Wen, B.-Z. Wan, Y.-M. Chang, C.-P. Chou, J.-M. Chen, W.-N. Chen, *Thin Solid Films* **515,** 7275 (2007).
7. H. Sinha, J.L. Lauer, M.T. Nichols, G.A. Antonelli, Y. Nishi, and J.L. Shohet, *Appl. Phys. Lett.* **96,** 052901 (2010).
8. M.-S. Kuo, G.S. Oehrlein, *J. Vac. Sci. Technol. B* **28,** 1104 (2010).
9. T. Yoshino, N. Ohnuki, N. Hata, Y. Seino, *Jpn. J. Appl. Phys.* **48,** 050210-1 (2009).
10. P.R. Aravind, P. Shajesh, G.D. Soraru, K.G.K. Warrier, *J. Sol-Gel Sci. Techn.* **54,** 105 (2010).
11. R.A. Orozco-Terana, B.P. Gormana, D.W. Mueller, M.R. Baklanovc, R.F. Reidy, *Thin Solid Films* **471,** 145 (2005).
12. G. Aksenov, D. De Roest, P. Verdonck, F.N. Dultsev, P. Marsik, D. Shamiryan, H. Arai, N. Takamure, M.R. Baklanov, *Mater. Res. Soc. Symp. Proc*. **1156,** D02-08 (2009).
13. I. Ciofi, M.R. Baklanov, Z. Tokei, G.P. Beyer, *Microelectron. Eng.* **87,** 2391 (2010).
14. J.L. Gurav, I.-K. Jung, H.-H. Park, E.S. Kang, D.Y. Nadargi, *J. Nanomat.* 409310, (2010).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.111-116

PACS numbers: 68.37.Hk, 42.70.Qs, 81.07.b, 81.16.Be

SYNTHESIS AND CHARACTERISATION OF $Cd_xZn_{1-x}S$ NANOCOMPOSITES

***M. Ragam*[1]*, K. Prem Ananth*[1]*, P. Parasakthi*[1]*, M. Thiribuvan*[1]*, K. Prema*[1]*, R Karuppaian*[1]*, N. Sankar*[1]*, K. Ramachandran*[2]**

[1] Department of Physics,
Yadava College, Govindarajan Campus,
Thiruppalai, Madurai 625 014, India
E-mail: ragamyc1983@yahoo.com

[2] School of Physics,
Madurai Kamaraj University,
Madurai-21, Tamilnadu, India

$Cd_xZn_{1-x}S$ nanoparticles have been synthesized using hydrothermal method. Structural characterization was done by XRD where the lattice structure gradually changes from hexagonal to cubic with increasing percentage of Zn in $Cd_xZn_{1-x}S$ nanoparticles. Optical spectroscopy provided evidence that the absorption edges of those nanoparticles can be varied from blue to UV. The nanoparticles exhibit emission peaks that shift to shorter wavelength with increasing percentage of Zn in the compounds $Cd_xZn_{1-x}S$. The control of the composition of $Cd_xZn_{1-x}S$ nanoparticles may lead the development of ideal materials for short wavelength diode laser applications.

***Keywords:** SEMICONDUCTORS, HYDROTHERMAL SYNTHESIS, OPTICAL PROPERTIES, SCANNING ELECTRON MICROSCOPY, CADMIUM ZINC COMPOSITES.*

(Received 04 February 2011, in final form 20 March 2011)

1. INTRODUCTION

The Synthesis and characterization of semiconductor nanoparticles have attracted much interest because of their novel properties as a consequence of the large number of surface atoms and the three dimensional confinement of the electrons [1]. Altering the size of the particles alters the degree of the confinement of the electrons and affects the electronic structure of the solid, especially the band gap edges, which are tunable with particle size. Among a variety of semiconductor materials, the binary metal chalcogenides of group II-IV have been extensively studied [2]. They have outstanding potential applications, owing to their nonlinear optical and luminescence properties [3], quantum size effect [4], and other important physical and chemical properties [5]. Nanocrystalline CdS and ZnS are attractive materials in photo conducting cells and optoelectronic devices such as solar cells and photodetectors [6]. Also the related ternary compounds $Cd_xZn_{1-x}S$ are promising materials for high density optical recording and for blue or even ultraviolet laser diodes.These applications are based on the quantum well structures of $Cd_xZn_{1-x}S$, which exhibits fundamental absorption edges that can be varied from green to UV [7].

to Contents, page 1

A facile, easy controlled and economical method was described in this article. We used simple inorganic raw materials to synthesize a series of nano- sized $Cd_xZn_{1-x}S$ samples involved in hydrothermal process. The central objective of the present study is to understand the growth behavior of ZnS-CdS nanoparticles, the optical and structural properties have been studied, and the results indicate that the absorption edges of those nanoparticles can be varied from blue to UV, while the emission peak can be adjusted from 470 to 690 nm which shows it has potential applications as wide band gap window materials.

2. EXPERIMENTAL

2.1 Chemicals

Chemicals were all of analytical reagent grade quality and used without further purification. Deionized and doubly distilled water were used throughout this study.

2.2 Synthesis

In a typical synthesis, 0.003 mol of $CdCl_2 \cdot 2H_2O$ and 0.003 mol of $Zn(AC)_2 \cdot 2H2O$ were dissolved in 41.67 mL of ethylene glycol to from a clear solution after stirring for 15 min at room temperature. 0.006 mol of $Na_2S.9H_2O$ were added into 20.34 mL of ethylene glycol which was then stirred to form a homogeneous solution. The above solutions were mixed and were then transferred into an autoclave with an inner Teflon lining and maintained at 140 ºC for 3 h. The yellow precipitate was harvested by centrifugation and washed several times with deionized water and ethanol to remove possible cations and anions before being dried in oven at 80 ºC for 6 h.

2.3 Characterization of $Cd_xZn_{1-x}S$ nanocomposites

Products were characterized by X-ray diffraction (XRD) recorded on a X-ray diffraction (PANalytical X'pert PRO X-Ray Diffractometer) with Cu k-α (λ = 1.54060 Å) as the source of incident radiation. UV-Vis spectra were recorded on a Shimadzu UV Spectrophotometer in the spectral range 200 - 900 nm at room temperature. The room temperature PL spectra are recorded using Perkin-Elmer LS-55 fluorescence spectrophotometer. Scanning Electron Microscopy (SEM) picture of these particles were taken using JEOL, JSM-840 microscope

3. RESULTS AND DISCUSSION

3.1 Crystal Structure and Morphology

Fig. 1. shows X-ray diffraction pattern of $Cd_xZn_{1-x}S$ nanoparticles. It is clear that XRD patterns of the nanoparticles exhibits prominent peaks at scattering angles (2θ) of 26.548, 43.37, and 51.91 which could be indexed to scattering from the (002), (110) and (112) planes respectively, of the hexagonal $Cd_xZn_{1-x}S$ crystal lattice. From the spectra it is clear that at lower Cd/Zn ratios, the crystal structure is similar to that of the cubic ZnS lattice. The (111) reflection is the main reflection observed from the ZnS sample, and this reflection is shifted to lower 2θ angle and the intensity of

to Contents, page 1

the reflection gradually decreases when the Cd concentration in the nanoparticles increases. This phenomenon was also observed by Laukaitis et al.

From the half width of the XRD peaks, the average particle size is tabulated in Table 1. based on Scherrer equation ($D = 0.9\lambda/B\cos\theta$), where D is the crystal diameter, λ is the X-ray wavelength 1.5408 Å and θ is the diffraction angle. But it is known that the accuracy of Scherrer formula for particle size below 5 nm is questionable as instrumental broadening of the diffractometer will mask. And so dislocation density and strain are calculated and tabulated in Table 1. These sizes are taken as only guideline values.

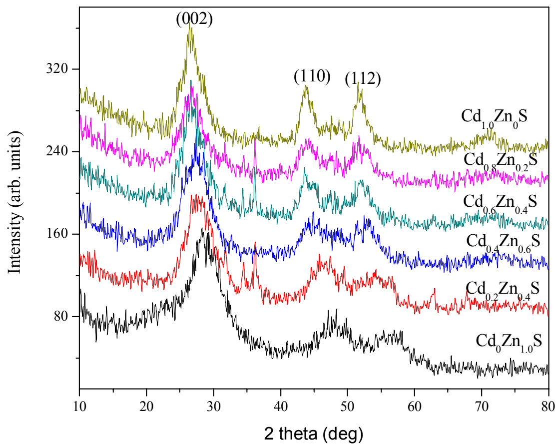


Fig. 1 *– XRD spectra for $Cd_xZn_{1-x}S$ (x = 0, 0.2, 0.4, 0.6, 0.8, 1.0) nanoparticles*

Table 1– *Structural parameters of $Cd_xZn_{1-x}S$ nanocomposites*

Samples	Particle size (nm)	Dislocation density (lines/m^2)	Strain
$Cd_0Zn_{1.0}S$	2.1	2.14×10^{17}	0.02
$Cd_{0.2}Zn_{0.8}S$	2.4	1.61×10^{17}	0.017
$Cd_{0.4}Zn_{0.6}S$	2.2	1.94×10^{17}	0.020
$Cd_{0.6}Zn_{0.4}S$	2.3	1.79×10^{17}	0.018
$Cd_{0.8}Zn_{0.2}S$	2.5	1.62×10^{17}	0.073
$Cd_{1.0}Zn_0S$	3.5	7.99×10^{16}	0.012

3.2 SEM Analysis

Fig. 2. shows the SEM images of $Cd_xZn_{1-x}S$ nanocrytals. Well-crystallized particles with estimated particle sizes of 30 - 40 nm were observed. The sample is composed of a large quantity of spherical nanoparticles with uniform size, shape and large specific surface area, it can significantly improve electron transfer rate which contribute to the high photo catalytic activity.

to Contents, page 1

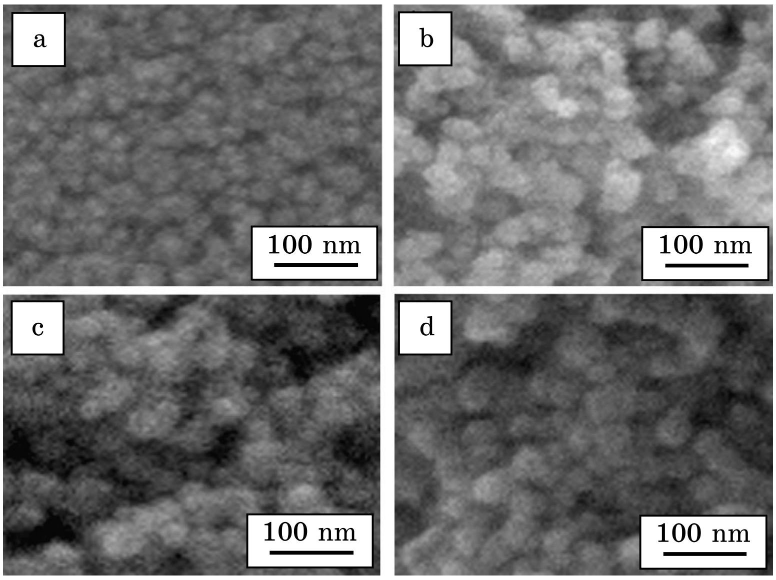


Fig. 2 – *SEM image of* $Cd_xZn_{1-x}S$ *(x = 0.2, 0.4, 0.6, 0.8) nanoparticles*

3.3 UV-Vis Analysis

The optoelectronics properties were investigated by ultraviolet-visible absorption spectra using the Shimadzu UV Spectrophotometer in the spectral range 200 - 900 nm at room temperature Fig. 3. shows the UV absorption spectra of $Cd_xZn_{1-x}S$ nanoparticles solutions along with those of the pure ZnS nanoparticles and CdS nanoparticles solutions. ZnS nanoparticles have an absorption peak at 302 nm, while CdS shows its absorption shoulder peak at 477 nm, the wavelengths corresponding the bandgaps of 4.11 eV for ZnS and 2.60 eV for CdS respectively. The blue shifts in the absorption edge bands have been interpreted in terms of quantum size effects [8] which are tabulated in table 2.

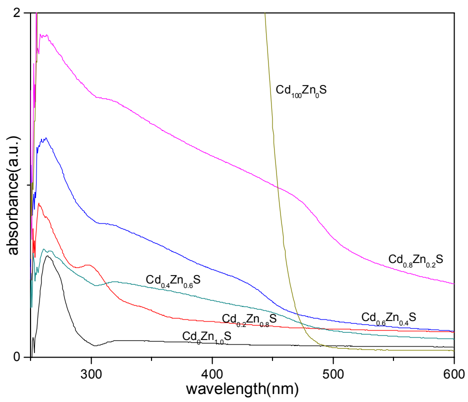


Fig. 3 – *UV-Vis spectra of* $Cd_xZn_{1-x}S$ *(x = 0.2, 0.4, 0.6, 0.8) nanoparticles*

to Contents, page 1

The absorption spectra of nanoparticles were found to lie between those of pure CdS and ZnS nanoparticles solutions. And the absoption edge bands of nanoparticles systematically redshifted from 302 nm to 450 nm with increasing the content of Cd in $Cd_xZn_{1-x}S$ nanoparticles which indicates the adsorption edges of those nanoparticles can be varied from blue to UV. This is very similar to the result obtained by Shengnan Zu et al. [8].

Table 2 – *Optical parameters of* $Cd_xZn_{1-x}S$ *nanocomposites*

Nanoparticles	Absorption edge (nm)	Bandgap (eV)
$Cd_0Zn_{1.0}S$	302	4.11
$Cd_{0.2}Zn_{0.8}S$	324,288	4.07
$Cd_{0.4}Zn_{0.6}S$	492,305	3.29
$Cd_{0.6}Zn_{0.4}S$	471,301	3.37
$Cd_{0.8}Zn_{0.2}S$	499,302	3.30
$Cd_{1.0}Zn_0S$	477	2.60

3.4 PL Analysis

The photoluminescence spectra of the $Cd_xZn_{1-x}S$ nanocrystals in Fig. 4 shows three peaks around 620 nm, 660 nm and 670 nm, lying between those of pure CdS and ZnS nanoparticles. The emission band around 480 nm for ZnS nanoparticles arises from the recombination through surface localized state [9]. For the distinct peak at 620 nm shows the electron hole recombination after relaxation, while emission around 660 nm is due to the trap state emission [10].

An interesting feature in these spectra is that once the composite nanoparticles form, the band edge emission of CdS nanoparticles will disappear. Therefore, the PL bands of $Cd_xZn_{1-x}S$ nanoparticles as in Fig. 4. are caused by trap state emission [10]. In our system, PL band position is red shifted as the Cd content increases, which is consistent with the results of UV absorption, confirming the formation of nanoparticles. And the position of emission peaks can be adjusted from 460 to 670 nm when increasing the Cd content in the $Cd_xZn_{1-x}S$ nanoparticles.

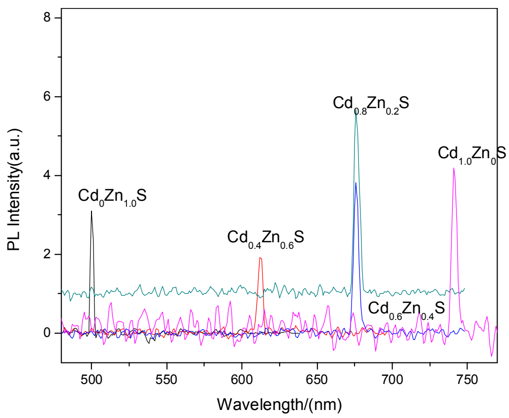


Fig. 4 – *PL spectra of* $Cd_xZn_{1-x}S$ (x = 0.2, 0.4, 0.6, 0.8) *nanoparticles*

to Contents, page 1

4. CONCLUSION

In conclusion, we have developed a novel and simple method by hydrothermal process to produce nano sized $Cd_xZn_{1-x}S$. The obtained nanocrystals present homogenous alloyed structure and uniform spheres. Lattice structure of $Cd_xZn_{1-x}S$ changes from cubic to hexagobal as x goes from 0 to 1. The optical properties of nano $Cd_xZn_{1-x}S$ can be modulated by tuning their compositions with a wide range which has the prospect for solar energy utilization due to the variable optical properties.

REFERENCES

1. A. Henglein, *Chem. Rev.* **89**, 1861 (1989).
2. M.L. Steigerwald, L.E. Brus, *Acc. Chem. Res.* 23, 183 (1990).
3. M.G. Bawendi, M.L. Steigerwald, L.E. Brus, *Annu. Rev. Phys. Chem.* 41, 477 (1990).
4. Y. Wang, N. Herron, *J. Phys. Chem.* 95, 525 (1991).
5. S. Gorer, J.A. Ganske, J.C. Hemminger, R.M.J. Penner, *Am. Chem. Soc.* **120**, 9584 (1998).
6. J. Britt, C. Ferekides, *Appl. Phys. Lett.* **62**, 2851 (1993).
7. T. Taguchi, Y. Endoh, Y. Nozue, *Appl. Phys. Lett.* **59**, 3434 (1991).
8. S. Zu, Z. Wang, B. Lie, X. Fan, G. Qian, *J. Alloy. Compd.* **476**, 689 (2009).
9. S. Wageh, S.M. Liu, T.Y. Fang, X.R. Xu, *J. Lumin.* **102**, 768 (2003)
10. Y. Li, F. Huang, Q. Zhang, *J. Mater. Sci.* **35**, 5933 (2000).
11. G. Laukaitis, S. Lindross, S. Tamulevicius, M. Leskela, M. Rackaitis, *Appl. Surf. Sci.* **396**, 161 (2000).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.117-126

PACS numbers: 61.05.C – , 68.37.Hk

STRUCTURE, OPTICAL AND ELECTRICAL CHARACTERIZATION OF TIN SELENIDE THIN FILMS DEPOSITED AT ROOM TEMPERATURE USING THERMAL EVAPORATION METHOD

***N. Kumar*[1], *V. Sharma*[1], *U. Parihar*[1], *R. Sachdeva*[1], *N. Padha*[1], *C.J. Panchal*[2]**

[1] Jammu University, Department of Physics and Electronics,
Ambadkar Road, 180006, Jammu, India
E-mail: nareshpadha@yahoo.com

[2] M.S. University of Baroda,
Faculty of Technology and Engineering, Vadodara, 390-00 India

Tin Selenide (SnSe) is an important IV-VI compound semiconducting material used for various devices like memory switching, an efficient solar cell and holographic recording systems. SnSe thin films of the thickness of 100 *nm were deposited by thermal evaporation method on a Glass substrate at room temperature. The prepared samples were investigated for structural, compositional, morphological and optical characterization respectively by using X-ray diffraction analysis (XRD), scanning electron microscopy (SEM) and transmission measurements. Thus deposited films showed a good polycrystalline quality having preferred (*111*) orientation with uniformly distributed spherical grains having size 16nm.The grown film identified as P- types by hot probe method. The films were found to have direct band transition having an optical bandgap (E_g) of* 1.92 *eV at room temperature. The temperature depended electrical resistivity (ρ) determined by using the two probe method, found to be* 390 Ω·*m at room temperature.*

Keywords: *TIN SELNIDE, THIN FILMS, SUBSTRATE TEMPERATURE, XRD, SEM, EDAX, TRANSMISSION ANALYSIS, ELECTRICAL RESISIVITY.*

(Received 04 February 2011)

1. INTRODUCTION

Metal chalcogenides offer wide range of optical band gaps suitable for various optical and optoelectronics applications. Among the IV-VI compounds, Tin Selenide (SnSe) has potential applications in memory switching devices, as solar cell material, in LASER and in holographic recording system. SnSe has an orthorhombic crystal structure and studied in the form of both single crystals and thin films [1-3]. Considerable efforts have been made by the researchers for the preparation of SnSe thin films by various techniques and this had led to several different methods viz. atomic layer deposition [4],chemical bath deposition [5], vacuum evaporation [6], chemical vapor deposition [7], spray pyrolysis [8], electrodeposition [9] and flash evaporation[10] etc. Literature survey on the subject matter reveals that the attempts were made by Subba Rao et al. [11], Bhatt et al. [12] Engelken et al. [13], Padiyan et at. [6] to obtain polycrystalline structure of the deposited films at room temperature but could not be succeeded and there reported temperature were found to 150 °C onwards at which the grown SnSe thin films were uniform, porous free and well adhesive with glass substrate. Tin Selenide (SnSe) thin films have been reported to grow at room temperature by using Thermal Evaporation technique.

to Contents, page 1

2. EXPERIMENT DETAILS

The SnSe thin films studied in this work were grown on a glass substrate at room temperature by using fine-grained pulverized SnSe powder (99.99 %), which was obtained from Alfa Aesar (USA) in a Hind Hivac Vacuum coating unit (model No 12-A4D). The vacuum was maintained at a base pressure of 10 mbar during deposition. The SnSe thin films considered in this study were deposited on soda lime glass substrates with dimensions of approximately $76 \times 25 \times 1$ mm^3. The quality of the substrate, prior to the growth of the thin films, is a crucial factor, which influences the material properties of the deposited thin films. Surface defects, such as scratches and dust on the substrate, have an adverse effect on the structural properties of the thin film. In order to obtain glass substrates with a high degree of chemical cleanliness, the following procedure of organic cleaning was used: 1. The glass substrate was rinsed in hydrogen peroxide to remove contaminants. 2. Substrate was then cleaned, in turn, under vapors of acetone, trichloroethylene, and methanol, respectively. The rate of deposition was 0.3 nm/s and typical thicknesses of the films were 100 nm that were continuously monitored during the deposition using a quartz crystal thickness monitor DTM-101 (Hindhivac, India).

2.1 Characterization Technique

The structural characterization of thin film under investigation was carried out using an X-ray diffractometer (XRD), D-Max-III (Rigaku), in 2θ range of 20°°- 60°, at a scan-rate of 0.05°s1, using Cu Kα (λ = 0.154 nm) radiation. The surface morphology of the films were studied using a Scanning Electron Microscope (SEM), JSM-5600 (JEOL), operated at 20 kV. To have an idea about the surface elemental composition of the film, Energy Dispersion Analysis by X-ray (EDAX) was carried out by using JEOL EDX Spectrometer (Model No.6360). The optical transmittance measurement was carried out with unpolarized light, at normal incidence, in the photon energy range of 0.8 - 2.5 eV, using the monochromator, CM110, photodetectors, and a lock-in amplifier, SR-530. The whole setup was automated using Lab View (Version 8.2). Temperature depended resistivity measurements for the samples were carried out in a Liquid Nitrogen bath in the temperature range 80 - 330 K using Keithley Model 6517A programmable electrometer. A Lake Shore model no 340-temperature controller is used for controlling and measuring the temperature (T).

3. RESULTS AND DISCUSSION

The films obtained as tested by visual inspection were found to be blackish gray in colour with good adhesion, which were under taken for structural, compositional, morphological optical and Electrical characterization. The structural analysis of the SnSe thin films has been carried out using X-ray Diffraction method with CuKα radiation. Fig. 1 shows the XRD Spectra of SnSe thin film of 100 nm thickness deposited on glass substrate at room temperature. The prominent Bragg reflection is occurring at or around 2θ = 30° corresponding to (111) diffraction plane, along with three other very weak diffraction peaks of (011), (311), (411), which confirms the polycrystalline nature of the film.

to Contents, page 1

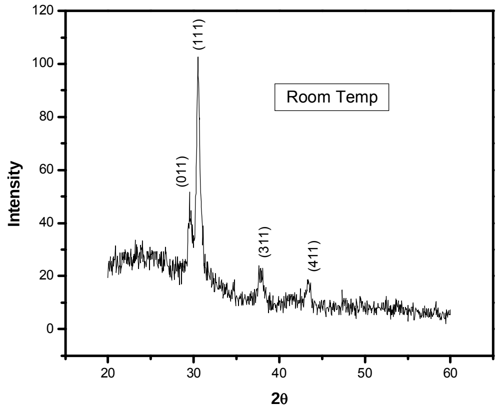


***Fig. 1** – XRD Spectra of SnSe thin film (*100 *nm) at room temperature*

A similar preferred orientation of (111) plane in SnSe film was observed by Bhatt et al. [12] and by Dang Tran Quan et al. [14] in the thin films grown by the Vacuum Evaporation Technique and by Singh & Bedi et. al. [15] prepared by Hot Well Epitaxy method. Whereas Teghil et al. [16] observed preferred orientation in the (011) and (200) crystallographic planes in the SnSe thin film prepared by Laser Ablation Method and John et al. [17] repeated (400) plane for films grown by Reactive Evaporation. The various preferred orientation reported for SnSe films indicate that the deposition technique plays an important role for the orientation of SnSe thin film deposition. The XRD data have also found useful for establishing dhkl, Crystallite size (D), Strain (ε) and Dislocation density (δ). The inter-planar spacing dhkl was calculated for the (111) plane using the Bragg's relation [18]

$$d_{hkl} = \frac{n\lambda}{2\sin\theta}, \tag{1}$$

where λ is the wavelength of the X-ray used, n is the order number and θ is Bragg's angle. The crystallite size (D) of the films was calculated from the Debye Scherrer's formula from the full width at half maximum (FWHM) of the peaks expressed in radians [18]

$$D = \frac{0.94\lambda}{\beta\cos\theta}, \tag{2}$$

where β is the FWHM calculated from the (111) plane. The dislocation density (ε) defined as the length of dislocation lines per unit volume of the crystal and calculated by using the formula

$$\delta = \frac{1}{D^2}. \tag{3}$$

The values of interplaner spacing (d) Grain size (D), and Dislocation density (δ) are calculated from Eq. (1), (2) and (3) are given in table no (1) respectively.

to Contents, page 1

Table 1 *– Structural parameter of SnSe film deposited on glass substrate at room temperature*

Substrate Temperature (°C)	Plane (hkl)	Thickness (E), nm	Interplanar Spacing(d), Å	Lattice const., Å	FWHM (Degrees)	Grain Size(D) (nm)	Dislocation density (1010 line\m^2)
Room Temperature	(111)	100	3.03807	5.25558	0.50742	16	0.05896

The analysis of the diffraction patterns also suggests that the SnSe thin-film deposited at room temperature has orthorhombic structure with lattice parameters $a = b = 0.431$ nm and $c = 0.540$ nm belonging to the D2h16 space group while the interplaner spacing d value corresponding to the (111) prominent peak is determined to be 3.03 nm which is in accordance with the d value given in the JCPDS data. Also, the position and intensities of the peaks are consistence with JCPDS card File (No. 32-1392). This shows that the film grown at room temperature has good crystallinity.

The strain (η), particle size (D) and dislocation density (δ) are also calculated by using the Williamson and Smallman relation (19)

$$\beta \cos\theta = \frac{\lambda}{D} - \eta \sin\theta \,, \tag{4}$$

where λ is the wavelength of the radiation used (0.15418 nm), β the full width at half maximum, and θ the angle of diffraction.

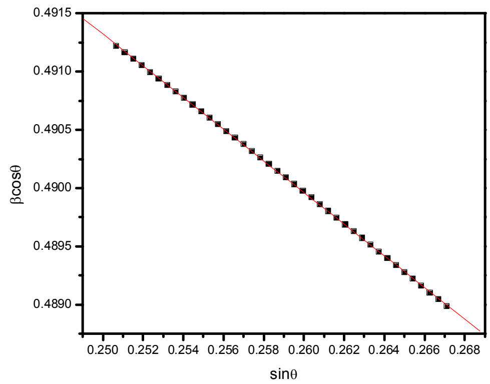


Fig. 2 *– Plot of βcosθ vs sinθ for a SnSe thin films of thickness of* 100 *nm grown at room temperature*

A graph is drawn between βcosθ and sinθ that provides a straight line as shown in figure (2). According to equation (4) the slope of graph provides strain (η) while particle size (D) is determined from the intercept. The value of grain size (D) and dislocation density (δ) obtain from the this method are same as that equation (3) obtained from using Scherrer's Method (20). The values of Strain (η), Grain size (D), and Dislocation density (δ) are calculated from eq. (4) are given in table no (2) respectively.

to Contents, page 1

Table 2 *– Micro-structural parameters of SnSe films deposited on glass substrate at room temperature*

Substrate Temperature °C	Plane (*hkl*)	Thickness (*E*), nm	Average internal strain (η)	FWHM (Degrees)	Average grain size, nm	Dislocation density (1010 line\m^2)
Room Temperature	(111)	100	0.13543	0.50742	16	0.05896

3.1 Composition analysis

The composition analysis of the SnSe thin films has been carried out using Energy dispersive analysis of X-ray (EDAX) recorded in the binding energy range 0 - 20 KeV. The EDAX spectra of SnSe thin film shown in Fig. 3 revealed that SnSe contents depend critically on the growth condition and technique used for deposition of thin films. From the EDAX patterns, the presence of Sn and Se peaks is observed and some other peaks are also observed which corresponds to Si, Na, Ca and O that can be attributed to the glass substrate used [21]. The atomic mass percentages SnSe of the films grown at room temperature have been found to be 31.78 and 33.18 respectively. This shows that the film grown at room temperature is slightly rich in Selenium and is nearly stoichiometric in nature, which is in agreement with the reported value of Tomkiewicz et al. (22) and Skyllas kazcos and Miller [23].

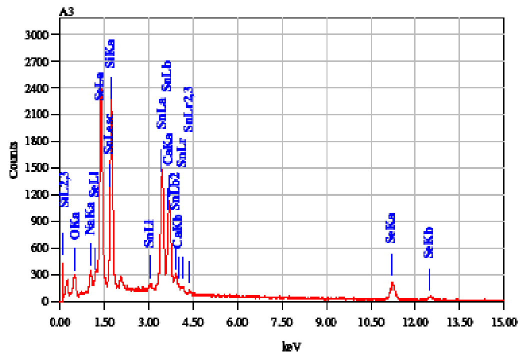


Fig. 3 *– EDAX Spectra of SnSe thin film (*100 *nm) deposited at room temperature on the glass substrate*

3.2 SEM Analysis

Scanning Electron Microscope(SEM) studies were carried out to assess the quality of the SnSe thin films deposited on glass substrate at room temperature. Fig. 4 shows a scanning Electron Micrograph of the synthesized SnSe thin films. The SEM micrograph shows a distribution of particles which covers the surface of the substrate completely. No pin holes or cracks could be observed for that sample. These results also confirm the results obtain from XRD data.

to Contents, page 1

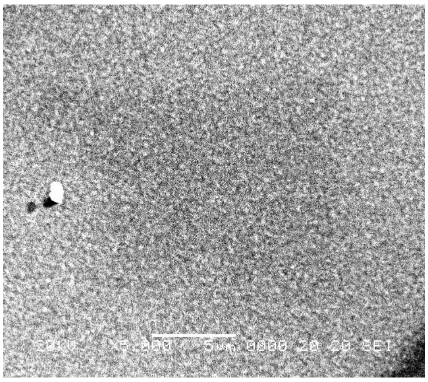


***Fig. 4** – Scanning Electron Micrograph of SnSe thin film (*100 *nm thick) deposited on the glass substrate at room temperature*

3.3 Optical Analysis

Optical studies are obtained by recording the transmission spectra of the thin film deposited on glass substrate in the wavelength range of 500 nm to 1500 nm at room temperature. Fig. 5 shows the optical transmission spectra of SnSe thin film deposited at room temperature. Film shows more than 50 % transmission for wavelength longer than 800 nm which is an indication of good crystallinity of the film.

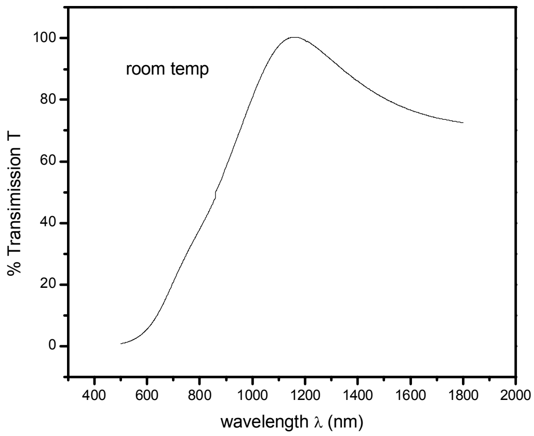


***Fig. 5** – Transmission spectra of SnSe thin film at room temperature*

The optical absorption coefficient was calculated using Lambert law (24)

$$\ln\left[\frac{I_0}{I}\right] = 2.303A = \alpha d \ , \qquad (5)$$

where I_0 and I are the intensity of incident and transmitted light, respectively; α is absorption coefficient, A is the optical absorbance and d is the film thickness.

to Contents, page 1

The spectral dependence of absorption coefficient i.e the function $a = f(hn)$ at room temperatures is shown in Fig. (6). The relation between absorption coefficient $(\alpha h\nu)^2$ and the photon energy $h\nu$ is given by the equation of Bardeen et al (25)

$$\alpha h\nu = B(h\nu - E_g)^x, \tag{6}$$

where B is the edge width parameter; $x = 0.5$ for direct transition, $x = 1.5$ for allowed and forbidden transition respectively, $x = 2$ for indirect allowed transition, $x = 3$ for indirect forbidden allowed transition. It is observed that the spectral variation can be described by equation (6) with $x = S$.

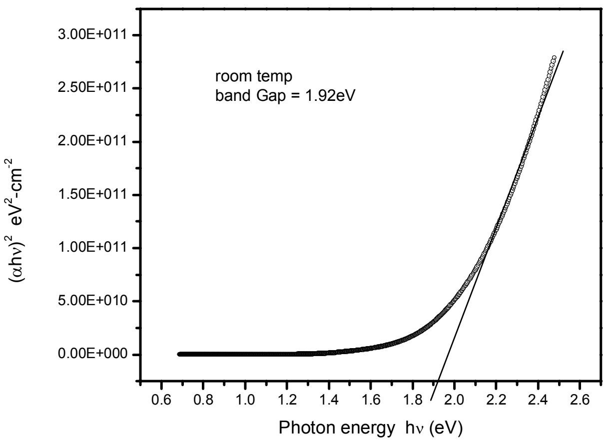


***Fig. 6** – Plot of $(\alpha h\nu)^2$ vs photon energy $h\nu$ of SnSe thin film (100 nm)*

The optical band gap E_g of the film was determine from the extrapolation of the linear fit plot of $(\alpha h\nu)^2$ versus at $h\nu = 0$ and was found to be 1.92 eV at room temperature. Linearity of above plot indicates that the material is of the direct band gap. The values of bands gap in agreement with the band gap values as repeated by other Nariya et al reported an indirect band gap value of 1.0 eV by direct vapor transport technique [26], Zulkarnain et al reported an indirect band gap of 1.25 eV by a combination of chemical precipitation and vacuum evaporation technique [27], N.A. Okereke et al. reported reported an indirect band gap of 1.5 eV by a combination of chemical deposition [31] and Matthew et al. reported the direct band gap value of 1.71 eV by a Solution-Phase Synthesis [32].

3.4 Electrical Analysis

The type of electrical conduction (p-type) in SnSe thin films was verified using the hot-probe method while Temperature depended electrical reisitivity measurement of semiconducting SnSe thin film was done by Two-probe method in the temperature ranges from 310 K down to 80 K. The room temperature electrical resistivity of the grown SnSe thin film was found to be 390 Ω·m.

to Contents, page 1

Fig. 5 shows the variation of electrical resistivity with change in temperature. The decrease in resistivity as temperature increase shows the semiconducting nature of the film [30].

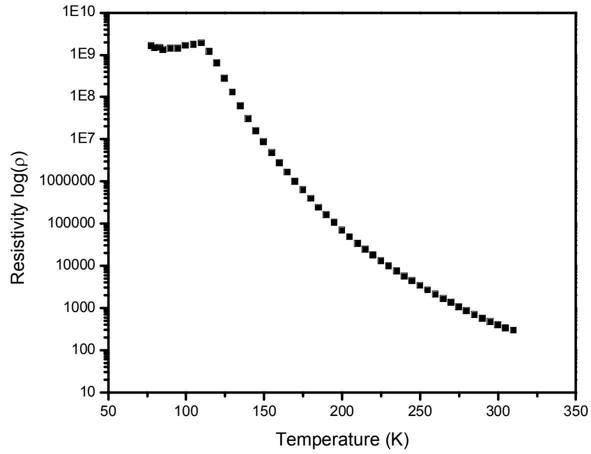


***Fig. 7** – Temperature depended Electrical Resistivity measurement of SnSe thin film using two probe method in the temperature range from* 325 *K down to* 80 *K*

The electrical resistivity of a polycrystalline thin film sample is a complex phenomenon, involving charge-carriers transport through both the "bulk-like" part of the semiconductor crystals and through the inter-crystalline (grain) boundaries. In the literature [28] the temperature dependence of the semiconductor material's resistivity is expressed by the equation no (7)

$$R = R_0 \exp\left[\frac{E_a}{K_b T}\right], \tag{7}$$

where R_0 is the pre-exponential factor, E_a is the activation energy for this thermally activated process and K_b is the Boltzmann constant.

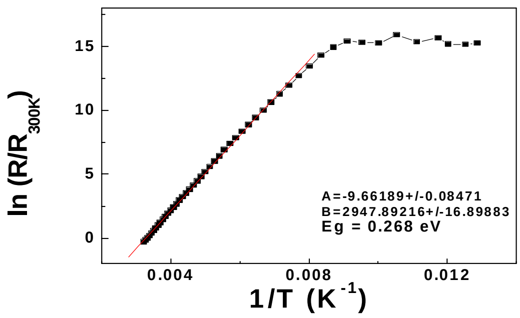


***Fig. 8** – Plot of $ln(R/R_{300K})$ versus $1/T$ for a SnSe thin film of thickness of* 100 *nm grown at* 350 *K temperature*

Clearly, a plot of $\ln(R/R_0)$ versus $1/T$ will be a straight line, from the slope of which the activation energy can be calculated. Thus, to measure the

to Contents, page 1

conductivity it is enough to measure the electrical resistance R since we are interested in the slope of the linear-least square fit only. So, the temperature dependence of resistance of SnSe thin films has been studied by measuring the resistance in the temperature range 80 - 330 K using the Keithley Model 6521 scanner card. Depending on the sample resistivity, a voltage limit was adjusted to obtain reliable data. The data collected was normally repeated for reproducibility check. After stabilizing to the desired temperature, the resistance values were normally recorded three times and their mean was noted. Once the dimensional factors were determined for each sample, the resistivity values were calculated. The resistivity thus obtained had an estimated error within 5%.The values of the activation energy for electrical conduction closely correspond with the measurements performed by another group [29].

4. CONCLUSION

Tin Selenide thin films have been successfully grown by thermal evaporation technique onto glass substrate held at 3050 K. X-ray Diffraction analysis confirmed that the deposited SnSe thin films were polycrystalline in nature having Orthorhombic structure with preferred orientation of grains along the (111) direction. Various structural parameters such as crystalline size, Strain and Dislocation Density are calculated from the XRD Spectra. SEM studies reveal that the SnSe films exhibited uniformly distributed grain over the entire surface of the substrate. The average sizes of the grains are found to be 16 nm. The presences of elemental composition were confirmed from EDAX analysis. Optical Transmittance measurements indicate the deposited film have a direct bandgap of 1.92 eV which confirms the formation of well crystallized SnSe films. Activation energy calculated from temperature dependent resistivity measurements was found to be 0.268 eV which correspond to shallow donor level near conduction band.

REFERENCES

1. A. Agarwal, S.H. Chaki, D. Lakshminarayan, *Mater. Lett.* **61**, 5188 (2007).
2. A. Agarwal, P.H. Trivedi, D. Lakshminarayan, *Cryst. Res. Technol.* **40**, 789 (2005).
3. Z. Zainal, S. Nagalingam, A. Kassim, M.Z. Hussian, W.M.M. Yunus, *Sol. Energy Mater. Sol. Cells* **81** No2, 261 (2004).
4. V.E. Drozd, I.O. Nikiforova, V.B. Bogevolnov, A.M. Yafyasov, E.O. Filatova, D. Papazoglou, *J. Phys. D: Appl. Phys* **42**, 125306 (2009).
5. Z. Zainal, N. Saravanan, K. Anuar, M.Z. Hussein, and W.M.M. Yunus, *Mater. Sci. Eng. B* **107**, 181 (2004).
6. D. P. Padiyan, A. Marikani, and K.R. Murali, *Cryst. Res. Technol.* **35**, 949 (2000).
7. N.D. Boscher, C.J. Carmalt, R.G. Palgrave, and I.P. Parkin, *Thin Solid Films* **516**, 4750 (2008).
8. L. Amalraj, M. Jayachandran, and C. Sanjeeviraja, *Mat. Res. Bull.* **39**, 2193 (2004).
9. Z. Zainal, A. Jimale Ali, A. Kassim, and M.Z. Hussein, *Sol. Energy Mater. Sol. Cells* **79** No2, 125 (2003).
10. G.H. Chandra, J. Naveen Kumar, N. Madusudhana Rao, S. Uthanna, *J. Cryst. Growth* **306**, 68 (2007).
11. T. Subba Rao, B.K. Samantharay and A.K. Chaudhuri, *J. Mat Sci Lett* **4**, 743 (1985).
12. V.P. Bhatt, K. Gireesan and C.F. Desai, *Crystal Res. Technol.* **24**, 187 (1989).

to Contents, page 1

13. R.D. Engelken, A.K. Berry, Doren Van, J.L.T.P. Boone and A.J. Shanazary, *Electro Chem. Soc.* **133**, 581 (1986).
14. D.T. Quan, *phys. status solidi a* **86**, 421 (1984).
15. J.P. Singh, R.K. Bedi, *J. Appl. Phys.* **68**, 2776 (1990).
16. R. Teghil, A. Santagata, V. Marotta, S. Orlando, G. Pizzella, A. Giardini-Guidoni, A. Mele, *Appl. Surf. Sci.* **90**, 505 (1995).
17. K.J. John, B. Pradeep, E. Mathal, *J. Matter. Sci.* **29**, 1581 (1994).
18. B.D. Cullity, *Elements of X-ray Diffraction* (Addison-Wesley: Reading, M.A.: 1972).
19. G.B. Williamson, R.C. Smallman, *Phil. Mag.* **1**, 34 (1956).
20. B.D. Cullity, *Elements of X-ray Diffraction* (Addison-Wesley Publ. Co.: California: 1956).
21. S. Prabahar, M. Bhanam, *J. Cryst. Growth* **285**, 41 (2005).
22. M. Skyllas Kazcos, B. Miller, *J. Electrochem. Soc.* **127**, 2016 (1980).
23. R.B. Kale, C.D. Lokhande, *Semicond. Sci. Tech.* **20**, 1 (2004).
24. A. Adachi, A. Kudo, T. Sakata, *Bull. Clm. Soc. Jpn.* **68**, 3283 (1995).
25. J. Bardeen, F.J. Blatt, L.H. Hall, R. Brechenridge, B. Russel, T. Hahn, *Photoconductivity Conf.* (Wiley: New York: 1956).
26. B.B. Nariya, A.K. Dasadia, M.K. Bhayan, A.J. Patel, A.R. Jani, *Chalcogenide Lett.* **6**, 549 (2009).
27. Z. Zulkarnain, J. Ali, K. Anuar, H. Mohd Zobir, *Malaysian Journal of Analytical Sciences* **7**, 197 (2001).
28. A.J. Moulson, *Electroceramics* (Wiley: USA: 1990).
29. K. Chung, D. Wamwangi, M. Woda, M. Wuttig and W. Bensch, *J Appl. Phys.* **103**, 083523 (2008).
30. N. Kumar, V. Sharma, N. Padha, N.M. Shah, M.S. Desai, C.J. Panchal, I.Yu. Protsenko, *Cryst. Res. Technol.* **45**, 53 (2010).
31. N.A. Okereke, A.J. Ekpunobi, *Chalcogenide Lett.* **7**, 531 (2010).
32. A. Franzman Matthew, W. Schlenker Cody, E. Thompson Mark, L. Brutchey Richard, *J. Am. Chem. Soc.* **132**, 4060 (2010).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.127-131

PACS numbers: 81.16.Nd, 68.37.Ps

NANO-SCALE PATTERNING OF SILICON NANOPARTICLES ON SILICON SUBSTRATE BY DIP-PEN-NANOLITHOGRAPHY

***A. Kumar*[1], *P.B. Agarwal*[1], *S. Sharma*[2], *D. Kumar*[2]**

[1] Central Electronics Engineering Research Institute,
Pilani – 333 031, India
E-mail: akumar1758@yahoo.co.in

[2] Department of Electronics Science, Kurukshetra University,
Kurukshetra – 136 119, India

Dip-Pen Nanolithography technique has been used to write nano-scale patterns of silicon nanoparticles on Si/SiO_2 substrate using commercially available silicon nanoparticles suspension as ink (mean diameter 30 nm). Patterning experiments have been carried out under varying process conditions namely, temperature and humidity with varying writing speed. Line- width of 92 nm has been measured at writing speed of 0.1 μm/sec, which reduced to 54 nm at higher speed of 1.6 μm/sec. Obtained results would be useful for patterning nano-size features of other hard materials (semiconductors and metals) for applications in nanoelectronics and biotechnology.

Keywords: *DIP-PEN NANOLITHOGRAPHY, SELF-ASSEMBLED-MONOLAYERS, AFM, SILICON NANOPARTICLES, NANOPATTERNING, NANOELECTRONICS.*

(Received 04 February 2011, in final form 18 March 2011)

1. INTRODUCTION

Continued innovations in process technology, materials, device-structures have made possible to develop 32 nm microprocessor chips. Patterning capability of lithography process determines the component density on the chips. Various lithographic techniques namely, electron-beam lithography [1], micro-contact printing [2], focused ion beam [3], scanning probe [4] and nanoimprint [5] are capable of delineating sub-50 nm patterns. Dip-pen-nanolithography (DPN) [6] is a new lithography tool primarily for writing molecules onto substrates through self-assembled-monolayers (SAM). Creating nano-structures using DPN is a single step process and does not require resist. Basic schematic of the technique is shown in Fig. 1.

In this technique, atomic force microscope (AFM) tip is used to write nano-dimensional pattern on the substrate. The AFM tip is coated with the desired compatible molecular/other ink. The ink molecules are transported to the substrate via a water meniscus which is formed between tip and substrate when in close proximity of substrate. The writing process strongly depends upon environmental conditions, such as humidity, temperature, tip coating procedures, surface condition and substrate-ink interactions. In case of thiol-based inks, chemisorption of ink-molecules on the underlying substrate act as a driving force for moving the molecules from the tip to the unoccupied sites on substrate. Various ink-substrate combinations have been explored for writing nano-pattern using dip-pen-nanolithography (DPN) [6].

to Contents, page 1

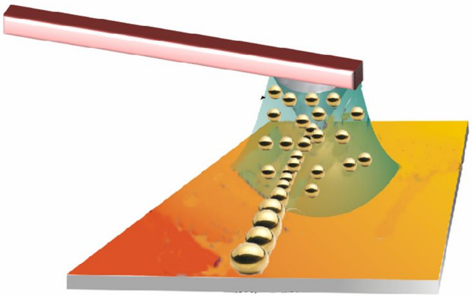


***Fig. 1** – Schematic representation of dip-pen-nanolithography (DPN) concept*

Recently, nano-patterning of metals, metal-oxides and semiconductors using dip-pen-nanolithography has been reported [9], [10]. In this work, the direct-writing capability of DPN has been explored in patterning of silicon nanoparticles on Si/SiO_2 substrate, using "suspension of silicon nanoparticles" as ink in place of generally used organic-molecule inks.

2. EXPERIMENTAL

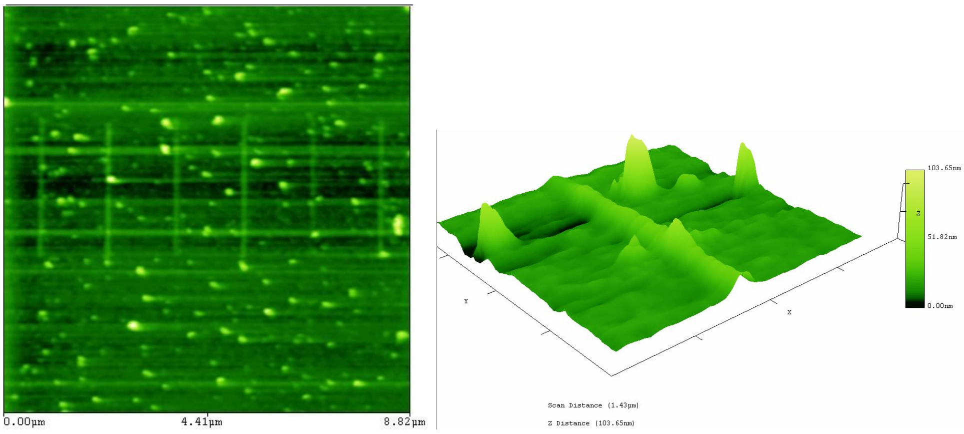
Low-resistivity, n-type, ⟨100⟩, silicon wafers were used as substrate in the experiments. Thermal oxide (~ 50 nm) was grown after chemical cleaning. Silicon nanoparticles (polycrystalline) suspension in IPA, purchased from Meliorum Technologies, USA, was used as writing ink. The specified mean diameter of nanoparticles is 30 nm with 10 % monodispersity. There are particles with diameter in range 20 - 44 nm, while 40 % particles have diameter in range 28 - 34 nm. A very dilute (5 mM) solution was prepared from the received suspension. Triangular shape silicon nitride cantilever AFM tips were coated with silicon nanoparticles suspension using dipping method. Writing experiments were conducted on NSCRIPTOR system from Nanoink, USA. First, several ink-diffusion test experiments were conducted under different temperature and humidity conditions in the environment-control chamber, to reach the favorable writing conditions. The reported experiments were performed at temperature = 27 °C and RH = 38 % at different locations on the substrate. Lines were designed in supplied software, InkCAD, and writting was performed with tip movement speeds in the range 0.1 - 1.6 μm/sec. Imaging of written lines was done using same AFM tip in lateral force microscopy (LFM) mode at higher frequency. Achieved AFM images have been analyzed and results are presented in the next section.

3. RESULTS AND DISCUSSION

LFM images, of lines written with tip movement speed from 0.1, 0.2, 0.4, 0.8, 1.2 and 1.6 μm/sec respectively are shown in Fig. 2a, and an enlarged 3-D view of one such line is shown in Fig. 2b. These images have been analyzed in image processing software module "NanoRule". For more accurate analysis, images of two lines were enlarged and line analysis was performed across a horizontal line drawn on image. Screen shot of such analysis is shown in Fig. 3. Two markers were used to measure line-width from the line profile as shown in the figure. The measured line widths

to Contents, page 1

varied from 92 nm to 54 nm by changing writing speed from 0.1 μm/sec to 1.6 μm/sec. The shown line widths are 65 nm and 54 nm for writing speeds of 1.2 μm/sec and 1.6 μm/sec respectively. Similarly, all six lines were processes and analyzed to estimate their width. The results are plotted in Fig. 4.



(a) (b)

Fig. 2 – LFM image lines of written using silicon nanoparticles ink with writing speeds of 0.1, 0.2, 0.4. 0.8, 1.2 *and* 1.6 *μm/sec respectively (a) and enlarged 3-D image of one such line (b)*

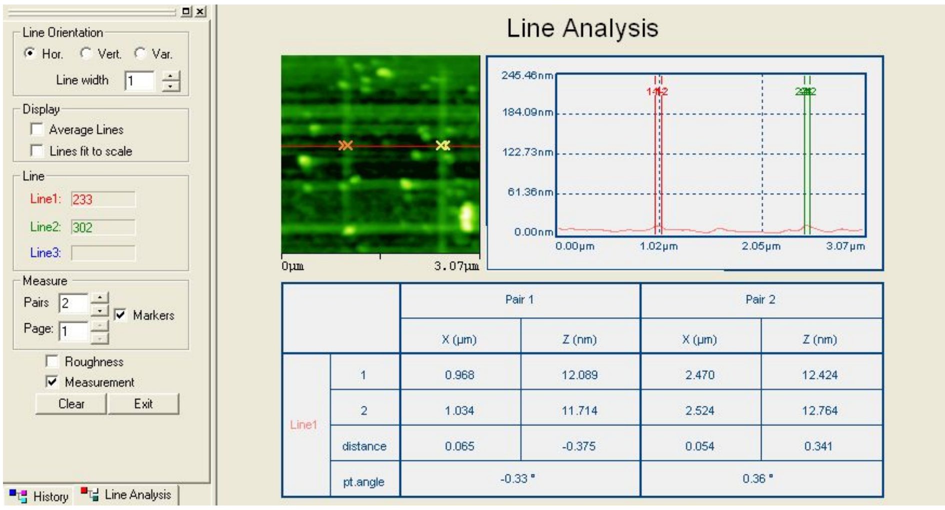


Fig. 3 *– Screen shot of image analysis of lines written with with speeds of* 1.2 *and* 1.6 *μm/sec, line profile across images and measured line widths are also shown*

Slow writing speed of 0.1 μm/sec resulted line of 92 nm width. Increasing speed to 0.2, 0.4, 0.8, 1.2 and finally 1.6 μm/sec, finer lines of width 80, 67, 67, 65 and 54 nm were achieved. Five sets of such experiments were performed on the same substrate and results were analyzed. Same trend of

to Contents, page 1

decreasing line width with increasing writing speeds was observed. This may be explained as follows; at lower speed, the tip is at given point for longer time, more silicon-nanoparticles are transported to the substrate and adsorbed on surface forming a line. With increasing speed, lesser number of particles are transported and deposited making finer line. There is little variation in line-width at these locations for a given speed. These small variations may due to topography and other variation on the surface condition of the substrate. As mentioned above, the mean diameter of nanoparticles is 30 nm, and there are particles of lower as well as higher size in the suspension. The written lines with measured width may be considered to contain 2-3 particles of different diameter aligned horizontally. Using highly mono-disperse nanoparticles suspension, lines with better control over size can be patterned having one or two particles across width.

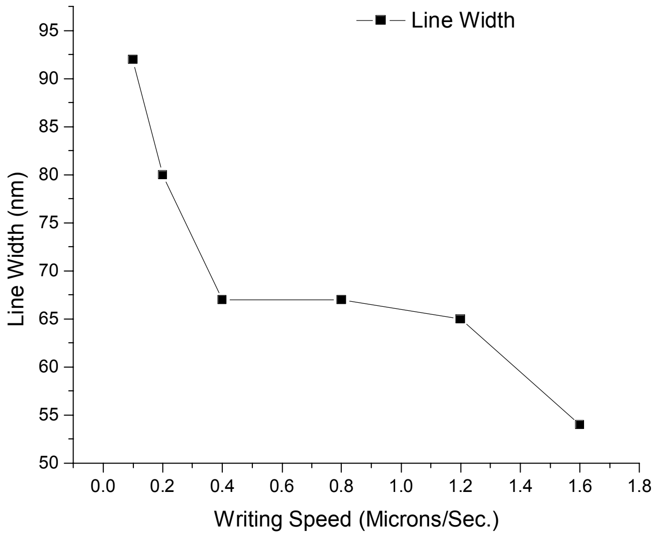


Fig. 4 *– Measured width of silicon nano-partilces lines with varying writing speeds from* 0.1 *to* 1.6 *μm/sec*

As the writing process is very sensitive to surface condition of local writing area, the lines have variation of $\pm$ 10% in their widths at different locations, with identical speeds. Few random bright spots are also visible in all the images. They appear to be nanometer size particles present on the substrate which might have been deposited during various fabrication processes of substrate. Great care is needed for substrate preparation and writing process. Surface contaminants and even roughness are detrimental to the controlled patterning. Storing the substrates after oxidation step in vacuum will improve the surface condition. Further, these bright spots may also be of some silicon nanoparticles diffused from the tip during movement and imaging. This problem may be controlled by reducing the amount of ink on the tip and imaging at higher speed. These lines may be considered as horizontal silicon nano-wires with silicon nanoparticles as beads, and find applications in nanoelectronic devices. These preliminary results demonstrate

to Contents, page 1

the direct nano-scale deposition of silicon nano-particles on silicon substrate at desired locations. Further studies of structure and electronic and optical properties are required before applying in real device. Similar experiment may be performed to write nano-size dots for applications in quantum dot/molecular devices.

4. CONCLUSIONS

Direct writing of silicon nanoparticles on silicon substrate has been demonstrated using Dip-pen-nanolithography technique. AFM imaging and analysis has been employed to determine line width. Minimum line width of 54 nm has been achieved at writing speed of 1.6 μm/sec., may be considered having two nano-particles across its width. Mono-disperse suspension is expected to result in more controlled features. Further process optimization in substrate fabrication, environment conditions and writing process are needed for patterning single nanoparticle-lines. These initial efforts would be useful in the area of nanoelectronics and nano-biotechnology.

ACKNOWLEDGEMENTS

Authors would like to thank Director, CEERI for his keen interest and support to this work, and Mr. R Sarvannan for image processing. This work has been carried out under R & D project funded by Department of Information Technology, Govt. of India. Financial support from DIT is gratefully acknowledged.

REFERENCES

1. C. Vieu, F. Carcenac, A. Pepin, Y. Chen, M. Mejias, A. Lebib, L. Manin-Ferlazzo, L. Couraud, H. Launois, *Appl. Surf. Sci.* **164,** 111 (2000).
2. J.A. Helmuth, H. Schmid, R. Stutz, A. Stemmer, H. Wolf, *J. Am. Chem. Soc.* **128,** 9296 (2006).
3. K. Arshak, M. Mihov, S. Nakahara, A. Arshak, D. McDonagh, *Superlattices Microstr.* **36,** 335 (2004).
4. A.A. Tseng, A. Notargiacomo, T.P. Chen, *J. Vac. Sci. Tech. B* **23,** 877 (2005).
5. W.J. Dausker, N.V. Le, E.S. Ainley, K.J. Nordquist, K.A. Gehoski, S.R. Young, J.H. Baker D. Convey, P.S. Mangat, *Microelectron. Eng.* **83,** 929 (2006).
6. D.S. Ginger, H. Zhang and C. Mirkin, *Angew Chem. Int. Ed.* **43,** 30 (2004).
7. B. Li, C. F. Goh, X. Zhou, G. Lu, H. Tantang, Y. Chen, C. Xue, F.Y.C. Boey, H. Zhang, *Adv. Mater.* **20,** 4873 (2008).
8. G. Lu, X. Zhou, H. Li, Z. Yin, B. Li, L. Huang, F. Boey, H. Zhang, *Langmuir* **26,** 6164 (2010).
9. H. Zhang, C.A. Mirkin, *Chem. Mater.* **16,** 1480 (2004).
10. B. Basnar, I. Willner, *Small* **5,** 28 (2009).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.132-139

PACS numbers: 61.72.Uj, 78.30.FS, 78.55.Et

SYNTHESIS AND CHARACTERIZATION OF ZnO NANOPARTICLES

***N. Singh*[1], *R.M. Mehra*[2], *A. Kapoor*[1]**

[1] Department of Electronic Science, University of Delhi south Campus, NewDelhi-110021, India
E-mail: Singhneetu1985@gmail.com

[2] School of Engineering and Technology, Sharda University, Greater Noida-201306 Uttar Pradesh, India

In this paper, we report the comparison between ZnO nanoparticles prepared via two different routes; i) via sol-gel route and ii) by solid state reaction method. It was found that when prepared under the same ambient conditions viz temperature, pressure etc. and keeping all the parameters same viz precursors, molarity, solvent etc; the nanoparticles prepared via Sol-gel route were highly crystalline and had smaller crystallite size (~ 24 nm) as compared to the one prepared by Solid state reaction method (~ 37 nm). The crystallinity and the crystallite size were examined by XRD and TEM. Variation in the bandgap as a function of size of the particles was determined using the absorption spectra obtained by UV-Vis-NIR spectrophotometer. Photoluminescence (PL) was also recorded in the visible region for the two types of particles and results have been analysed.

Keywords: *SYNTHESIS, ZnO, NANOPARTICLES.*

(Received 04 February 2011, in final form 17 March 2011)

1. INTRODUCTION

Zinc Oxide is a member of II-VI semiconducting compounds and occurs naturally as the mineral zincite. It is a hexagonal wurtzite type crystal exhibiting anisotropy. ZnO is a well-known n-type semiconductor and has got a wide band gap of 3.3 eV at 300 K. Some of its important properties are listed in Table 1 [10-16]. ZnO is considered a good candidate for transparent conducting electrodes in solar cells because it is transparent to the visible light [1]. It is also considered as a prime candidate for UV and blue light-emitting devices such as blue LED and Lasers due to its large exciton binding energy of 60 meV [2, 3]. Due to such large exciton binding energy, the excitons remain dominant in optical processes even at room temperature. Due to its vast industrial applications such as electro-photography, electroluminescence phosphorus, pigment in paints, flux in ceramic glazes, filler for rubber products, coatings for paper, sunscreens, medicines and cosmetics, ZnO is attracting considerable attention in powder as well as thin film form. Its resistance to radiation damages also makes it useful for space applications. The fabrication of ZnO nanostructures have attracted intensive research interests [4-5] as these materials have found uses as transparent conducting oxides (TCO) [8-9]. Since it is the hardest of the II-VI family of semiconductors with a large shear modulus, its performance is not degraded as easily as the other compounds through the

to Contents, page 1

appearance of defects. Since Zinc, the main constituent is cheap, non-toxic and abundant, ZnO has become commercially viable.

In the present work, we have synthesized ZnO nanoparticles via two different routes (sol-gel route and solid state reaction method) and tried to analyze the two on the basis of their crystallinity, crystallite size, bandgap and structural properties. X-ray diffraction (XRD) is used to calculate crystallite size. Variation in the bandgap as a function of size of the particles is determined using the absorption spectra obtained by UV-Vis-NIR spectrophotometer. Photoluminescence (PL) is also recorded in the visible region. Transmission electron micrograph (TEM) and Scanning electron micrograph (SEM) images are shown to clearly see the particle size and grain size respectively.

Table 1 *– Properties of zinc oxide*

Properties	Values
Crystal structure	Rock salt, Zinc blende and Wurtzite
Energy Bandgap, eV	3.2-3.3
Electron Mobility, cm^2Vs^{-1}	2.5-300 (Bulk ZnO), 1000 (Single nanowire)
Exciton Binding Energy, meV	60
Density, g/cm^3	5.606
Refractive Index	2.0041
Electron Effective Mass (m_e)	0.26
Relative Dielectric Constant	8.5
Melting point, °C	1975
Boiling point, °C	2360
Electron Diffusion Coefficient, cm^2s^{-1}	5.2 (Bulk ZnO), 1.7×10^{-7} (Particulate Film)

2. EXPERIMENTAL

2.1 Sol-gel route

All the reagents used were of analytical grade and no further purification was done before use. ZnO nanopowder was prepared by dissolving 0.2M Zinc acetate dihydrate [$Zn(CH_3COO)_2.2H_2O$] in methanol at room temperature and then mixing this solution ultrasonically at 25 °C for 2h. Clear and transparent sol with no precipitate and turbidity was obtained. 0.02 M of NaOH was then added in the sol and stirred ultrasonically for 60 min. The sol was kept undisturbed till white precipitates were seen in the sol. After precipitation, the precipitates were filtered and washed with the excess methanol to remove starting material. Precipitates were than dried at 80 °C for 15 min on hot plate.

2.2 Solid state reaction method

The chemical reagents used in this work were Zinc acetate dihydrate [$Zn(CH_3COO)_2 \cdot 2H_2O$] and NaOH powders of analytical grade purity. In solid-state reaction method, 0.2 M of Zinc acetate dihydrate [$Zn(CH_3COO)_2 \cdot 2H_2O$] was ground for 10 min and then mixed with 0.02 M of NaOH. After the mixture was ground for 30 min, the product was washed many times with

to Contents, page 1

deionized water and methanol to remove the by-products. The final product was then filtered and dried into solid powders at 80 °C for 15 min on hot plate.

The XRD measurements were carried out using Bruker AXS-D8 discover diffractometer. The absorbance of the powder in the visible region was measured using Shimadzu UV-VIS-NIR spectrophotometer (solidspec-3700). PL measurements were carried out using Shimadzu RF-5301 PC spectrofluorophotometer under 325 nm excitation wavelength. The crystallite size calculated from the XRD measurements was confirmed by TEM Morgagni-268D FEI.

3. RESULTS AND DISCUSSION

The XRD patterns of the nanoparticles obtained by sol-gel route and solid state reaction method are shown in Fig. 1 and 2 respectively. The nanoparticles synthesized by both methods showed crystalline nature with 2θ peaks lying at 31.750° ⟨100⟩, 34.440° ⟨002⟩, 36.252° ⟨101⟩, 47.543° ⟨102⟩, 56.555° ⟨110⟩, 62.870° ⟨103⟩, 66.388° ⟨200⟩, 67.917° ⟨112⟩, 69.057° ⟨201⟩, 72.610° ⟨004⟩, 76.95° ⟨202⟩, 81.405° ⟨104⟩, and 89.630° ⟨203⟩. The preferred orientation corresponding to the plane ⟨101⟩ is observed in both the samples. These peak positions coincide with JCPDS card no. 36-1451 for ZnO powder. Crystallite size was obtained by Debye-Scherrer formula given by equation

$$D = \frac{0.94\lambda}{\beta\cos\theta}, \tag{1}$$

where D is the crystallite size, 0.94 is the particle shape factor which depends on the shape of the particles, λ is the CuK_α radiations (1.54 Å), β is full width at half maximum (FWHM) of the selected diffraction peak corresponding to 101 plane and θ is the Bragg angle obtained from 2θ value corresponding to maximum intensity peak in XRD pattern. The crystallite size obtained using this formula is 23.585 nm for sol-gel derived particles and 37.344 nm in case of solid state reaction method derived particles.

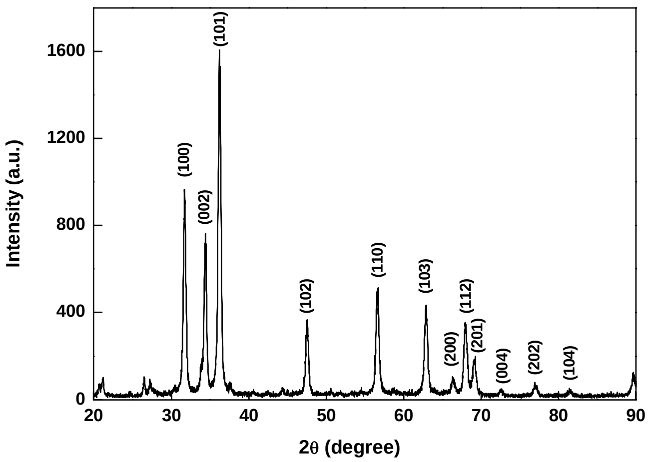


Fig. 1 *– XRD pattern of ZnO nanoparticles synthesized via sol-gel route*

to Contents, page 1

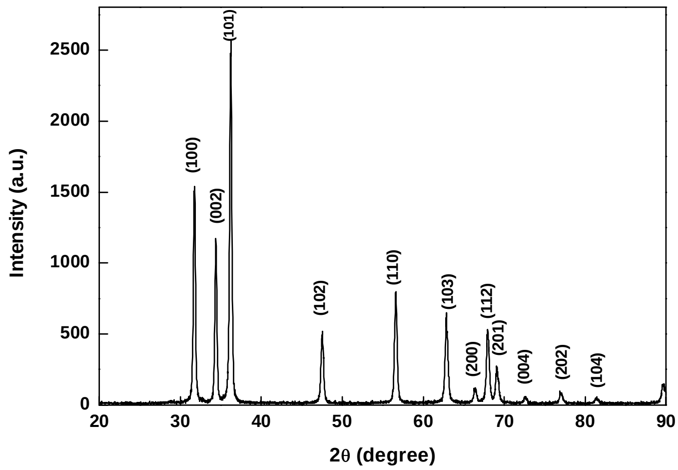


Fig. 2 *– XRD pattern of ZnO nanoparticles synthesized via solid state reaction method*

The absorbance curve of the sol-gel derived nanoparticles in the visible region is shown in Fig. 3(a). The graph shows that ZnO does not absorb light in the visible region. This result is in accordance with the bandgap value of the bulk ZnO (3.37 eV) according to which ZnO absorbs light in ultra violet (UV) range. Band gap energy is calculated using Tauc's plot Fig. 3b which comes out to be 3.23 eV in case of sol-gel derived nanoparticles. Tauc's equation (2) is given by [17]

$$\alpha h\nu = A(h\nu - E_g)^n, \quad (2)$$

where α is the absorption coefficient, $h\nu$ is the photon energy, A is the constant, E_g is the bandgap energy of the sample. The value of n is 1/2 or 2 depending upon whether the transition from valence band to conduction band is direct or indirect. The value is 1/2 in case of direct transition and 2 in case of indirect transition. Since ZnO is a has a direct band structure, the value of n is 1/2 in this case. So, the equation takes the form

$$(\alpha h\nu)^2 = B(h\nu - E_g), \quad (3)$$

where, B is a constant related effective masses of charge carriers associated with valence and conduction bands. Intersection of the slope of $(\alpha h\nu)^2$ vs $h\nu$ curve provides bandgap energy of the samples. According to the experimentally calculated bandgap, the synthesized ZnO nanoparticles should absorb light below 383 nm and absorbance graph is in agreement with this.

The absorbance curve of the solid state reaction derived nanoparticles in the visible region is shown in Fig. 4a. Tauc's plot is shown in Fig. 4b. The band gap energy comes out to be 3.15 eV from the Tauc's plot. According to the experimentally calculated bandgap, the synthesized ZnO nanoparticles should absorb light below 393 nm in this case and absorbance graph shows this thing. The band gap values validates our crystallite size results according to which smaller crystallite size should have larger band gap (23.585 nm, 3.23 eV for sol-gel derived nanoparticles) and large crystallite size should have smaller band gap (37.344 nm, 3.15 eV for solid state reaction derived nanoparticles).

to Contents, page 1

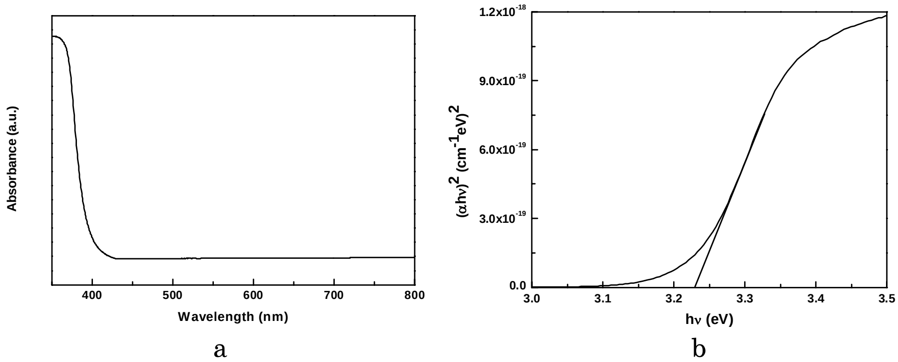


Fig. 3 *– Absorbance of sol-gel derived nanoparicles visible range (a) Tauc's plot of sol-gel derived nanoparticles (b)*

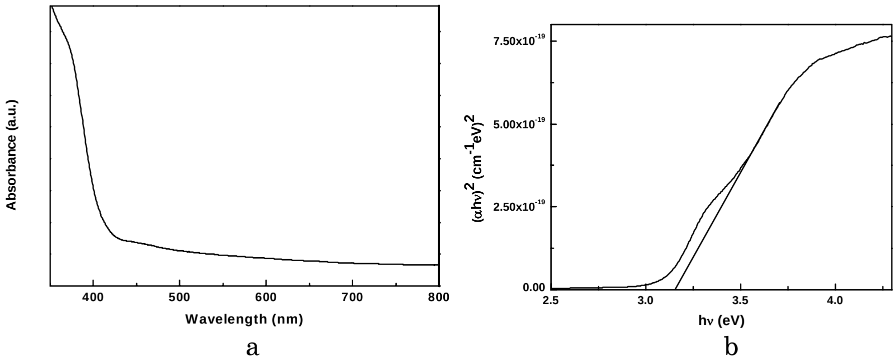


Fig. 4 *– Absorbance of solid state reaction derived nanoparticles in visible rang (a) Tauc's plot of solid state reaction derived nanoparticles (b)*

Photoluminescence (PL) spectra of the nanoparticles obtained by both the processes are shown in Fig. 5. The first peak in PL spectra corresponds to band to band transition and the spectra between 420-500 nm are showing blue luminescence. ZnO nanoparticles prepared via solid state reaction method show high luminescence than sol-gel derived nanoparticles. This could be due to the chemical instability caused during the fabrication process. As can be seen from the PL spectrum of sol-gel derived nanoparticles, the intensity peak is observed at 388.6 nm. If we calculate the band gap value from this wavelength, it comes out to be 3.2 eV. The PL intensity peak in case of solid state reaction derived nanoparticles is observed at 391.5 nm. From this value, band gap comes out to be 3.16 eV. The band gap energies calculated using PL spectra are approximately same as the ones calculated using Tauc's plot.

TEM images of sol-gel derived nanoparticles are shown in Fig. 6a. Clear hexagonal structures can be seen in the Fig. 6c having diameter ~ 23 nm. Selected area diffraction is shown in Fig. 6b which clearly indicates that the ZnO nanoparticles are highly crystalline in nature.

to Contents, page 1

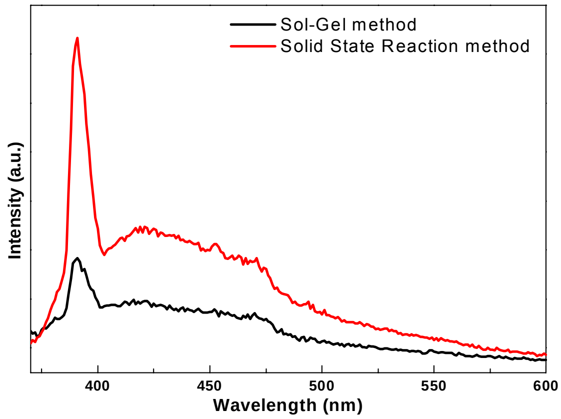


Fig. 5 *– Photoluminescence peak of ZnO nanoparticles obtained via different method*

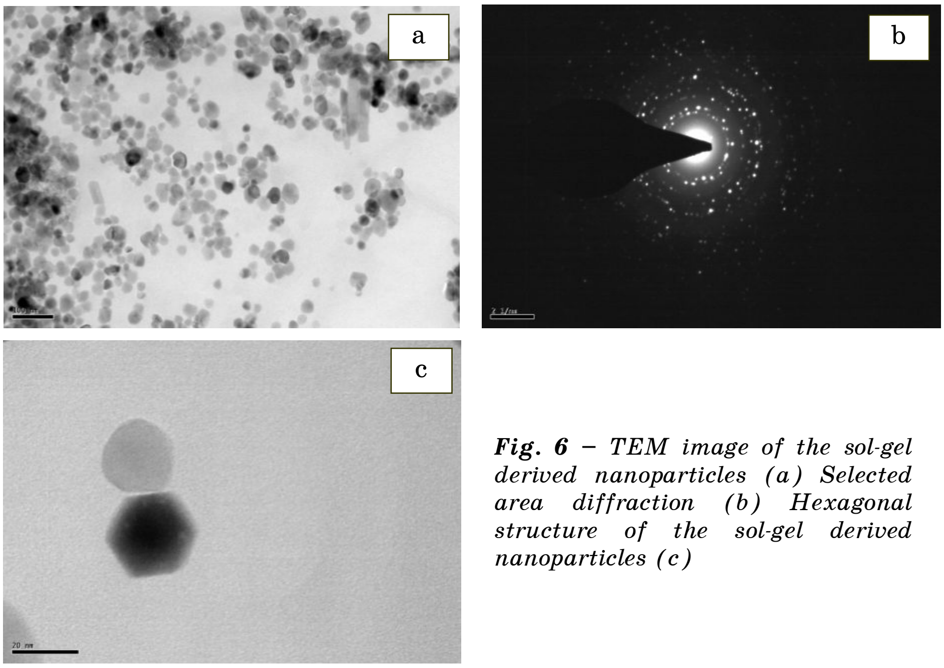


Fig. 6 *– TEM image of the sol-gel derived nanoparticles (a) Selected area diffraction (b) Hexagonal structure of the sol-gel derived nanoparticles (c)*

TEM image and selected area diffraction pattern of the solid state reaction derived nanoparticles are shown in Fig. 7a and 7b respectively. Selected area diffraction pattern of the nanoparticles indicates that the ZnO nanoparticles prepared via solid state reaction method are crystalline in nature. However the diffraction rings in this case are not properly aligned as in the case of sol-gel derived nanoparticles. No clear hexagonal structures can be seen in the TEM image. Nanoparticles obtained in this case are adhering to one another. Agglomeration of nanoparticles is more in this case than the former one. As can be seen from the TEM image that the average particle size is ~ 37 nm which is in agreement with the crystallite size obtained from XRD.

to Contents, page 1

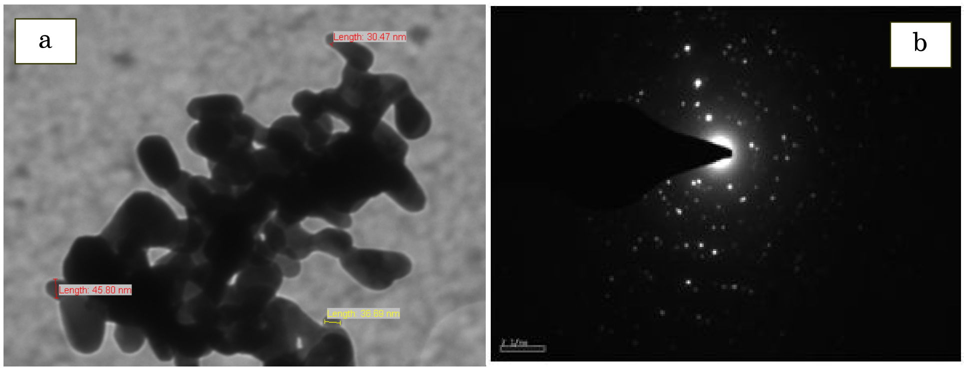


***Fig. 7** – TEM image (a) Selected area diffraction of solid state reaction derived ZnO nanoparticles (b)*

SEM images of the nanoparticles prepared via both the routes are shown in Fig. 8. Fig. 8a shows the SEM image of sol-gel derived nanoparticles. Clear nanostructures can be seen having grain size of ~ 70 nm. The crystallite size as observed from TEM in this case is ~ 24 nm. This shows that one grain in sol-gel derived nanoparticles is approximately equal to three crystallites. So it is clear that the nanoparticles seen by SEM image consist of a number of crystallites which are seen by TEM image. SEM image of nanoparticles prepared by solid state reaction method is shown in Fig. 8b. Grain size in this case is ~ 200 nm. Crystallite size as seen from TEM image is ~ 37 nm in this case. This shows that one grain in solid state reaction derived nanoparticles consists of approximately five crystallites. XRD results are confirmed by the combined study of these SEM and TEM images.

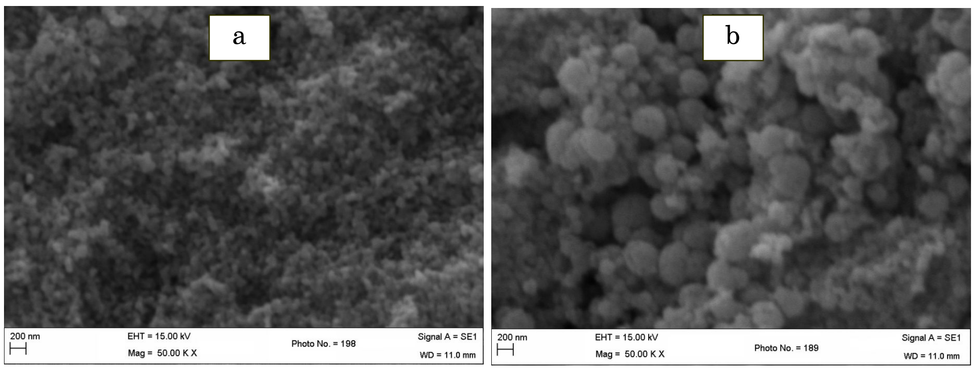


***Fig. 8** – Scanning electron micrographs via sol-gel route (a) via solid state reaction method (b)*

4. CONCLUSION

ZnO nanoparticles were prepared via sol-gel and solid state reaction methods. The ZnO nanoparticles prepared via sol-gel route were highly crystalline and had smaller crystallite size (~ 24 nm) as compared to the one prepared by Solid state reaction method (~ 37 nm). The bandgap of the synthesised nanoparticles was found to be size dependent. Photoluminescence (PL) study confirms the results obtained by XRD and TEM.

to Contents, page 1

ACKNOWLEDGMENTS

This work is done under the DST-JSPS sponsored project DST/INT/JSPS/PROJ/10. Financial support by University Grant Commission, India in the form of JRF is also acknowledged. The authors are thankful to University Science Instrumentation Centre (USIC) for providing XRD facility, All India Institute of Medical Sciences (AIIMS) for providing TEM facility and Prof. P.K. Bhatnagar from Department of Electronic Science, Delhi University for providing PL facility.

REFERENCES

1. D.R. Lide, *Hand Book of Chemistry and Physics, 71st ed.* (CRC, Boca Raton, FL: 1991)
2. A. Ohtomo, M.Kawasaki, I. Ohkubo, H. Koinuma, T. Yasuda, Y. Segawa, *Appl. Phy. Lett*. **75**, 980 (1999).
3. Z.K. Tang, G.K.L. Wong, P. Yu, M. Kawasaki, A. Ohtomo, H. Koinuma, Y. Segawa, *Appl. Phys. Lett.* **72**, 3270 (1998).
4. Y.W. Wang, L.D. Zhang, G.Z. Wang, X.S. Peng, Z.Q. Chu, C.H. Liang, *J. Crystal Growth* **234**, 171 (2002).
5. M.H. Huang, S. Mao, H. Feick, H. Yen, Y. Wu, H. Kind, E. Weber, R. Russo, P. Yang, *Science* 292 1897 (2001).
6. W. Zhu, C. Bower, O. Zhou, G. Kochanski, S. Jin, *Appl. Phys. Lett.* **75**, 873 (1999).
7. C.J. Lee, S.C. Lyu, Y. Zhang, H. Ruh, H.J. Lee, *Appl. Phys. Lett.* **81**, 3648 (2002).
8. J.W. Bae, S.W. Lee, K.H. Song, J.I. Park, J.J. Park, Y.W. Ko, G.Y. Yeom, *Jpn. J. Appl. Phys.* **38**, 2917 (1999).
9. G.K. Paul, S. Bandopadhya, S.K. Sen, *phys. stat. solid a* **191** No2, 509 (2002).
10. Ozgur, Y.I. Alivov, C. Liu, A. Teke, M.A. Reshchikov, S. Dogan, V. Avrutin, S.J. Cho, H. Morkoc, *J. Appl. Phys.* **98**, 41301 (2005).
11. H. Tang, K. Prasad, R. Sanijines, P.E. Schmid, F. Levy, *J. Appl. Phys.* **75**, 2042 (1994).
12. G. Oskam, Z.S. Hu, R.L. Penn, N. Pesika, P.C. Searson, *Phys. Rev. E* **66**, 11403 (2002).
13. H.S. Bae, M.H. Yoon, J.H. Kim, S. Im, *Appl. Phys. Lett.* **83**, 5313 (2003).
14. K. Keis, A.Roos, *Opt. Mater.* **20**, 35 (2002).
15. V. Noack, H. Weller, A. Evchmuller, *J. Phys. Chem. B* **106**, 8514 (2002).
16. J.J. Wu, G.R. Chen, C.C. Lu, W.T. Wu, J.S. Chen, *Nanotechnology* **19**, 105702 (2008).
17. A.B. Kashyout, M. Soliman, M.El Gamal, M. Fathy, *Mat. Chem. Phys*. **90**, 230 (2005).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.140-145

PACS number: 62.23.Hj

FABRICATION OF $Zn_xCd_{1-x}Se$ NANOWIRES BY CVD PROCESS AND PHOTOLUMINESCENCE STUDIES

***R.P. Vijayalakshmi*[1], *G. Murali*[1], *D. Amaranatha Reddy*[1], *R. Venugopal*[2]**

[1] S.V. University,
Dept. of Physics, 517 502, Tirupati, A.P. India
E-mail: vijayaraguru@gmail.com

[2] Govt. Degree College, Puttur.
Dept. of Physics, Chittoor, A.P. India

$Zn_xCd_{1-x}Se$ alloy nanowires with composition $x = 0.2$, 0.5 have been successfully synthesized by a simple thermal evaporation on the silicon substrate coated with a gold film of 20 Å thickness. The as-synthesized alloy nanowires, 70 - 150 nm in diameter and several tens of micrometer in length. The nanowires are single crystalline revealed from Transmission electron microscopy (TEM) and XRD measurement. The structure of $Zn_xCd_{1-x}Se$ nanowires are hexagonal wurtzite with [01-10] growth direction. Energy gap of the $Zn_xCd_{1-x}Se$ nanowires are determined from micro photoluminescence measurements. The energy gap increases with increasing Zn concentration.

Keywords: *$Zn_xCd_{1-x}Se$ NANOWIRES, CVD, PHOTOLUMINESCENCE, SEMICONDUCTOR NANOWIRES, II-VI NANOWIRES, TEM, XRD.*

(Received 04 February 2011, in final form 20 March 2011)

1. INTRODUCTION

The ternary semiconductor compounds $Zn_xCd_{1-x}Se$ due to their excellent optical properties and fast response times [1, 2] have a wide range of potential applications, especially in optical switching, optical communications, optical signal processing and photo voltaic devices [3, 4].

The processing advantages of II-VI materials compared with other semiconductors are their high photochemical stability and size-dependent optical properties due to confinement effect [5, 6]. Aside from choosing various semiconducting materials of different band gaps, it is now possible to control the band gap energy of a given semiconductor from lowering the dimensionality and or reducing its size to values comparable with or smaller than the corresponding excitonic Bohr diameter [5]. Alloying of semiconductors is another means that can be applied to achieve semiconductor materials for various band gap energies. Among all groups, II-VI materials are a good representative example in laser photonics to generate coherent blue-green light. The recently developed ZnSe based laser diodes have a potential coverage of the whole green spectral region (490 - 590 nm). The $Zn_xCd_{1-x}Se$ alloy is used as a quantum well materials in ZnSe based diodes grown on InP substrate which can effectively reduce the compressive strain.

to Contents, page 1

Alloyed nanowires of ternary II-VI semiconductors with elementary compositions of $Cd_xZn_{1-x}Se$ and ZnCdS [7, 8], successfully fabricated by MBE, MOCVD and laser assisted deposition [9]. Chang et al. [10] synthesized $Zn_xCd_{1-x}Se$ nanowires by MBE technique to photodetector applications. ZnSe/ZnCdSe heterostructures were grown on Si substrate [11] for luminescence applications.

2. EXPERIMENTAL STUDIES

Many methods have been developed to synthesize one dimensional nanostructures of II-VI semiconductors. Among these are thermal evaporation [12-14], laser ablation [16-17], arc discharge [18] and chemical synthesis methods [19]. The most common deposition based synthesis methods are physical and chemical vapor deposition [20-22]. Both methods consist of physical transport of the vapor species to the deposition site. To prepare $Zn_xCd_{1-x}Se$ ($x = 0.2$ and 0.5) nanowires a simple vacuum tube furnace is used (Fig. 1).

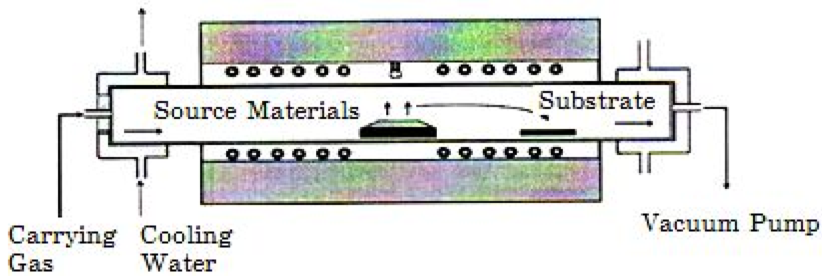


Fig. 1 *– Schematic of Single-zone tube furnace vapor-solid growth*

For the preparation of source material high purity ZnSe and CdSe (99.999 purity) were used. It was prepared by physical mixing the desired quantities of ($x = 0.2$ and 0.5) of ZnSe and CdSe and then sintering the mixture at 1000 ºC in a vacuum sealed quartz tubes for 18 hours. The sintered mixture was slowly cooled to room temperature in 5 hours and then used as a source material for the fabrication of $Zn_xCd_{1-x}Se$ alloy nanowires. The source materials are put in the alumina boat and placed in the centre of a single – zone horizontal tube furnace where the atmosphere, evaporation time, pressure and temperature are controlled. The Au film (20 Å thickness) coated on a Si substrate (20 × 10 mm) was used as a product collection substrate and located down stream in a lower temperature region in the furnace which kept at a reasonable vacuum. The substrate was vacuum annealed at 550 °C to recrystallize into Au nanoparticles. Then the temperature in the furnace is elevated to a controlled temperature of 900 °C at a specific rate. A carrier gas mixture of Ar (90 %), and H_2 (10 %) with flow rate of 220 sccm was frequently introduced into the quartz tube till the pressure was 250 torr. The furnace temperature was kept at 900 °C throughout the experiment.

The nanowires were characterized by XRD, SEM, TEM and Photoluminescence studies. X-ray diffraction (XRD) spectra of the as-synthesized nanowires were recorded using scintag X1 diffractometer with

to Contents, page 1

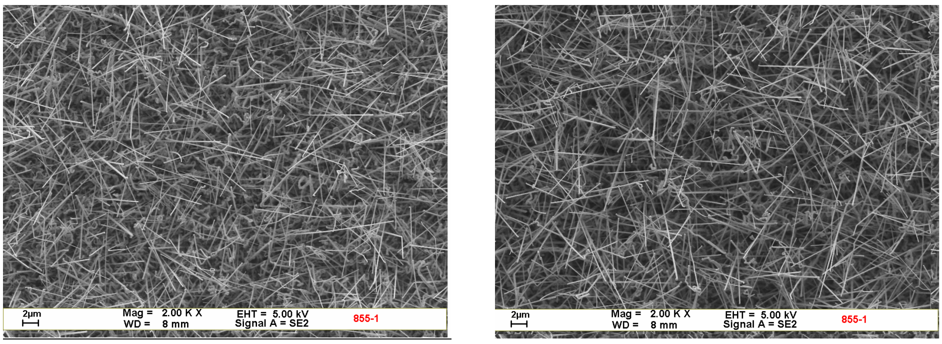
CuK_{α} (λ = 1.5418 Å) radiation at scanning speed of 2°/min in 2θ ranging from 20° to 60°. Composition were measured by energy dispersive X-ray spectroscopy (EDS). TEM was recorded from JEOL JEM 2010 analytical at 200 kV. Room temperature photoluminescence was measured in a confocal microscope (Jobin Yvou, MFO) using 325 Ar-Kr ion laser (coherent, Innova 70 °C) as an excitation source with laser intensity of $\leq$ 3.8 kW/cm^2.

2.1 Growth Mechanism

In this study the formation of $Zn_xCd_{1-x}Se$ (x = 0.2 and 0.5) nanowires follow a vapour-liquid-solid (VLS) growth mechanism as evidenced by the eutectic tips containing Au, Zn, Cd and Se. At an early reaction stage in the formation of $Zn_xCd_{1-x}Se$ nanowires, a vapour mixture of Zn, Cd and Se was carried by the Ar and H_2 gases and deposited onto catalytic Au nanodroplets. When the dissolution of Zn, Cd and Se in the Au nanodroplets became supersaturated, $Zn_xCd_{1-x}Se$ nanowires extruded from the liquid eutectic Au nanodroplets and precipitated at the liquid-solid interface. This process complies basically with an ordinary VLS growth model proposed originally by wagner et al. [23], in which a liquid cluster of metal catalyst provides energetically favored sites for the absorption/adsorption of gas-phase reactants. The sizes of the catalysts are considered to be responsible for the diameter of resultant nanowires. Growth directions of the $Zn_xCd_{1-x}Se$ nanowires are [01-10]. These results suggest that the $Zn_xCd_{1-x}Se$ alloy nanowires prefer a specific crystalline structure with a particular growth direction at certain composition ratio and growth temperature.

3. RESULTS AND DISCUSSION

General SEM morphologies of the as synthesized $Zn_xCd_{1-x}Se$ (x = 0.2 and 0.5) nanowires orienting randomly on the Si substrate are shown in figures 2a and 2b respectively. The diameter of the $Zn_xCd_{1-x}Se$ nanowires are distributed in the range of 70 - 150 nm. While the lengths of the nanowires are several tens of micrometer and respond sensitively to the synthetic reaction time. Chemical contents of the $Zn_xCd_{1-x}Se$ nanowires were analyzed from EDS spectra and their data shown in Figures 3a and 3b being listed in Table 1.



a b

Fig. 2 *– SEM image of $Zn_xCd_{1-x}Se$ nanowires (a) x = 0.2, (b) x = 0.5*

to Contents, page 1

***Table 1** – Chemical contents in the $Zn_xCd_{1-x}Se$ nanowires with various compositions*

Composition (x)	Chemical contents from EDS		
	Zn	Cd	Se
0.2	9.95	39.80	50.25
0.5	25.25	25.37	49.37

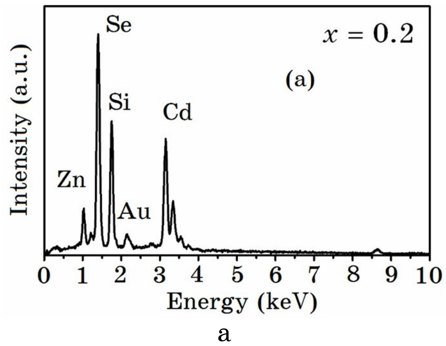


***Fig. 3** – EDS spectra of $Zn_xCd_{1-x}Se$ nanowires (a) $x = 0.2$ (b) $x = 0.5$*

Fig. 4 displays typical XRD pattern for the $Zn_xCd_{1-x}Se$ nanowires with $x = 0.2$ and 0.5 respectively. From the figure it is obvious that for both the compositions the structure is hexagonal wurtzite. The XRD spectra of nanowires of CdSe and ZnSe, CdSe and ZnSe nanopowders were recorded to compare the lattice parameters. We have calculated $a/c = 4.271/6.898$ Å (CdSe nanowires) and 4.293/6.998 Å for (CdSe powders), $a = 5.646$ (ZnSe nanowires) and 5.679 (ZnSe powder) for the zinc blende reveal that there exist lattice contractions in both of the as synthesized CdSe and ZnSe nanowires.

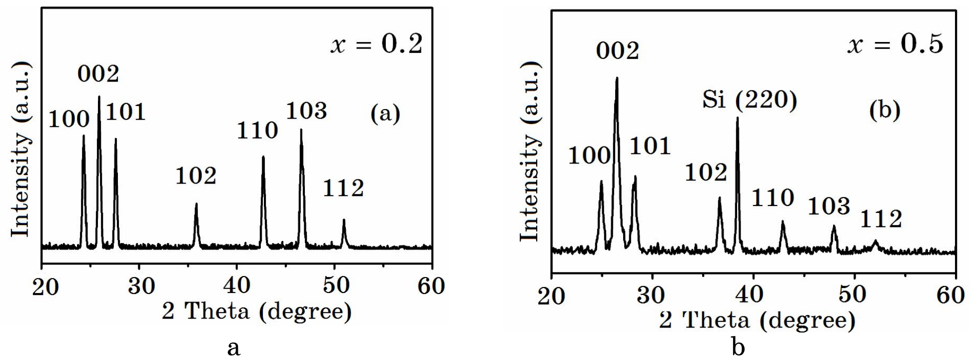


***Fig. 4** – XRD pattern of $Zn_xCd_{1-x}Se$ nanowires (a) $x = 0.2$ (b) $x = 0.5$*

to Contents, page 1

a b

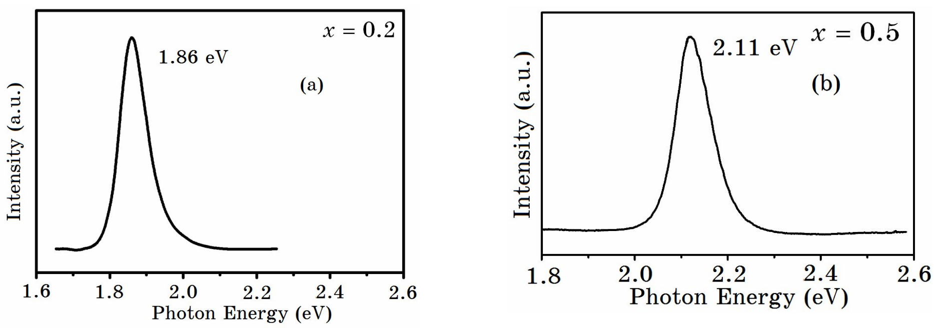
***Fig. 5** – TEM image of single nanowire of $Zn_xCd_{1-x}Se$ (a) $x = 0.2$ (b) $x = 0.5$*

The observed lattice contractions could have been induced by a surface tension along surface reconstruction in the growth of nanocrystallites, similar type of results were observed in CdSe nanocrystals [24] and CdSe nanobelts/nanosheets [25].

Typical TEM images of the $Zn_xCd_{1-x}Se$ ($x = 0.2$ and 0.5) are shown in Fig. 5a and 5b respectively. In any as-synthesized $Zn_xCd_{1-x}Se$ nanowires on Au-containing tip, composed of major Au and minor Zn, Cd and Se, was always found at one end suggesting that the nanowires compiled with a vapor-liquid-solid growth mechanisms.

3.1 Photoluminescence and energy gap

ZnSe and CdSe are direct band gap semiconductors. As synthesized $Zn_xCd_{1-x}Se$ alloy nanowires exhibit strong PL at room temperature. Fig. 6a and b shows the PL spectrum of $Zn_xCd_{1-x}Se$ nanowires. It is obvious from the figure that as Zn concentration increases the luminescence peak position shifted to higher energy, it is shown in Fig. 6a and b. The band gap was calculated from the strong peak and the band gap increases with Zn concentration.



a b

***Fig. 6** – Room temperature PL spectra of $Zn_xCd_{1-x}Se$ nanowires(a) $x = 0.2$ and (b) $x = 0.5$*

to Contents, page 1

4. CONCLUSIONS

$Zn_xCd_{1-x}Se$ alloy nanowires have been fabricated by CVD method. XRD studies confirmed that the crystal structure is hexagonal wurtzite. The synthesized nanowires exhibited photoluminescence, with increasing Zn concentration luminescence peak shifted to higher energy side. The band gap also calculated from luminescence peak. Band gap increased with Zn concentration.

Authors are grateful to thank University Grants Commission, New Delhi for providing financial support through Major Research Project.

REFERENCES

1. S.S. Yao, C. Kara guleff, A. Gabel, R. Fortenberry, C.T. Seaton, G.I. Stegeman, *Appl. Phys. Lett.* **46**, 801 (1985).
2. J. Yumoto, S. Fukuskima, K. Kubodera, *Opt. Lett.* **12** , 832 (1987).
3. R.S. Feigelson, A.N. Diaye, S.Y. Yin, R.H. Bube, *J. Appl. Phys* **48**, 3162 (1977).
4. S.I. Najafi, M. Belanger, R. Maliejko, R. J. Black, *Appl. Opt.* **27**, 806 (1988).
5. L.E. Brus, *J. Chem. Phys* **80**, 4403 (1984).
6. S.V. Ivanov, *Phys. Status Solidi A* **192**, 157 (2002).
7. A. Colli, S. Hofmann, A.C. Ferrari, F. Martelli, S. Rubini, C. Ducati, A. Franciosi, J. Robertson, *Nanotechnology* **16**, S139 (2005).
8. C.X. Shan, Z. Liu, C.M.Ng, S.K. Hark, *Appl. Phys. Lett.* **87**, 033108 (2005).
9. Y. Liu, J.A. Zapien, Y.Y. Shan, C.Y. Geng, C.S. Lee, S.T. Lee, *Adv. Mater.* **17**, 1372 (2005).
10. S.J. Chang, C.H. Hsiao, B.W. Lan, S.C. Hung, B.R. Huang, S.J. Young, Y.C. Cheng, S.H. Chih, *Superlattices Microst.* **48**, 50 (2010).
11. C.H. Hsiao, S.J. Chang, S.C. Hung, Y.C. Cheng, B.R. Huang, S.B. Wang, B.W. Lan, S.H. Chih, *J. Crystal Growth.* **312**, 1670 (2010).
12. X.M. Meng, J. Liu, Y. Jiang, W.W. Chen, C.S. Lee, I. Bello, S.T. Lee, *Chem. Phys. Lett.* **382**, 434 (2003).
13. Y. Jiang, X. Meng, J. Liu, Z. Xie, C.S. Lee, S.T. Lee, *Adv. Mater.* **15**, 323 (2003).
14. Q. Li, C. Wang, *Appl. Phys. Lett.* **83**, 359 (2003).
15. Y. Jiang, X.M. Meng, W.C. Yiu, J. Liu, J.X. Ding, C.S. Lee, S.T. Lee, *J. Phys. Chem. B* **108**, 2784 (2004).
16. P. Lefebvre, T. Richard, J. Allegre, H. Mathieu, A. Combette Roos, W. Granier, *Phys. Rev. B.* **53**, 15440 (1996).
17. X.Y. Kong, Z.L. Wang, *Nano Lett.* **3**, 1625 (2003)
18. Y.J. Xiong, Y. Xie, Z.Q. Li, X.X. Li, S.M. Gao, *Chem. Eur. J.* **10**, 654 (2004).
19. B. Gilbert, F. Huang, Z. Lin, C. Goodell, H. Zhang, J. Banfield, *Nano Lett.* **6**, 605 (2006).
20. S.S. Garje, J.S. Ritch, D.J. Eisler, M. Afzaal, P.O. Brien, T. Chivers, *J. Mater. Chem.* **16**, 966 (2006).
21. P.V. Radovanovic, C.J. Barrelet, S. Gradecak, F. Qian, C. Lieber, *Nano Lett.* **5**, 1407 (2005).
22. J.E. Mahan, *Physical Vapor Deposition on Thin Films*, (Wiley-Interscience, New York, NY: 2000).
23. R.S. Wagner, W.C. Ellis, *Appl. Phys. Lett.* **4**, 89 (1964).
24. Y.Z. Zhang, X.Y. Young, M. Xiao, *Appl. Phys. Lett.* **81**, 2076 (2002).
25. R. Venugopal, P.I. Lin, C.C. Liu, Y.T. Chen, *J. Am. Chem. Soc.* **127**, 11262 (2005).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.146-150

PACS numbers: 81.15.Pq, 61.46.Km

TEMPLATE ASSISTED GROWTH OF ZINC OXIDE-BASED NANOWIRES BY ELECTROCHEMICAL DEPOSITION

T. Singh, D.K. Pandya, R. Singh

Thin Film Laboratory, Department of Physics,
Indian Institute of Technology Delhi, Hauz Khas, New Delhi-110016, India
E-mail: triloks.iitd@gmail.com

Ordered ZnO and $Zn_{1-x}Cd_xO$ nanowire/nanorod arrays were fabricated by cathodic electrodeposition based on anodic alumina (AAO) membrane and polycarbonate membrane (PCM) from an aqueous solution containing zinc nitrate precursor at different bath temperatures. The electrodeposition process involves the electroreduction of nitrate ions to alter the local pH within the pores and precipitation of the metal oxide within the pores. X-Ray diffraction measurements showed that the nanowires/nanorods were of wurtzite crystallographic structures and the average length and diameter of nanorods were measured by SEM and TEM. HRTEM measurements confirm the crystallinity and elemental composition of grown nanowires on PCM/AAO templates.

Keywords: *ZnO, NANOWIRES, ELECTRODEPOSITION, TEM.*

(Received 04 February 2011, in final form 18 March 2011)

1. INTRODUCTION

Template-assisted synthesis has been employed widely to prepare solids of defined dimension [1]. Several techniques of forming materials in templates have been developed, including chemical vapor deposition(CVD) [2], sol-gel deposition [3], in situ polymerization [4] and electrodeposition [5]. Nanostructured materials have attracted great interest due to their unique chemical and physical properties, which can be influenced not only by the preparation procedure but also by their shape and size [6, 7]. The morphology of the nanostructures plays a key role especially on the optoelectronic properties of the materials, which determine the performance of semiconductors to be used in solar cells, as photo transistors and diodes, transparent electrodes, and so on. Among group II – VI semiconductor materials, ZnO is one of the most attractive functional semiconductor material for the fabrication of optoelectronic devices operating in the blue and ultraviolet region because of a direct wide band-gap of 3.37 eV and an exciton binding energy of 60 meV [8, 9]. It is well known that the realization of bandgap engineering to create barrier layers and quantum wells in device heterostructure is an important step for the design of ZnO-based devices. CdO is an n-type semiconductor with a direct band-gap of 2.3 eV and an indirect band-gap of 1.36 eV [10]. The ternary $Zn_{1-x}Cd_xO$ alloy can allow the bandgap tuning from 3.37 eV (band-gap of ZnO) to a narrower band-gap, i.e., into the visible spectral range, and their preparation have been reported [11]. The control over morphology and size of semiconductor materials represents a great challenge in realizing the

to Contents, page 1

design of novel functional devices. In this paper, we investigated an electrodeposition method for the growth of ZnO-based nanowires/nanorods by applying a negative potential to the substrate. Electrochemical deposition allows mixing of the chemicals at atomic level thus reducing the possibility of undetectable impurity phases, [12] and it is a good candidate to solve the problem of the small thermodynamic solubility of CdO in ZnO [13] Furthermore, the electrochemical deposition presents a simple, quick and economical method for the preparation of $Zn_{1-x}Cd_xO$ nanorods.

2. EXERIMENTAL

ZnO and ZnCdO nanowires/nanorods were synthesized by electrodeposition into nuclear track etch polycarbonate membrane (PCM) and anodic alumina membrane (AAO) (Whatman). The rated membrane thickness, nominal pore size, and pore density of PCM and AAO membranes were 6 μm, 100 nm, 6×10^8 pores/cm^2 and 60 μm, 100 nm, 8×10^8 (pores/cm^2) respectively. One side of the membrane was first coated with a 100 nm thick layer of gold, to serve as working electrode. The electrical contact was made to the membrane working electrode using gold coated silicon substrate. A platinum sheet 2×3 cm^2 was used as counter electrode and saturated calomel electrode was used as reference electrode. Electrosynthesis was performed under potentiostatic control using a Perkin Elmer 260 A instrument. The electrolyte (bath) temperature was varied from 70 to 90 °C. Electrosynthesis of ZnO was in an electrolyte containing zinc nitrate precursor solution. For the growth of ternary compound nanowires/nanorods the precursor solutions were obtained by varying the zinc nitrate and cadmium nitrate concentrations in de-ionized water. The cadmium concentration for the growth of nanostructures was varied from 4 to 16 atomic percent. The electrochemical deposition of ternary ZnCdO compounds were carried out at the deposition potential of -1.0 V (vs. SCE) 20 - 40 min. After the deposition the sample was removed from electrolyte and rinsed in de-ionized water. For the structural studies, X-ray diffractometer (Philips Xpert Pro) using CuK_α ($\lambda = 1.5405$ Å) radiation in 2θ range 20 - 800 was used. Scanning electron microscopy images were obtained using EVO-50. Energy dispersive X-ray spectroscopy analysis was obtained from Bruker-ASX (QuanTax 200).

3. RESULTS AND DISCUSSION

The electrodeposition process of ZnO nanowires/nanorods mainly includes two parts: first, an increase of pH, in this case due to the reduction of nitrate ions, and second the precipitation of zinc oxide. On reduction of nitrate in the presence of zinc ions the following two reactions occurs:

$$NO_3^- + H_2O + 2e^- \rightarrow NO_2^- + 2OH^- \quad (1)$$

$$Zn^{2+} + 2OH^- \rightarrow ZnO + H_2O \quad (2)$$

The total reaction may be written as,

$$Zn^{2+} + NO_3^- + 2e^- \rightarrow ZnO + NO_2^- \quad (3)$$

to Contents, page 1

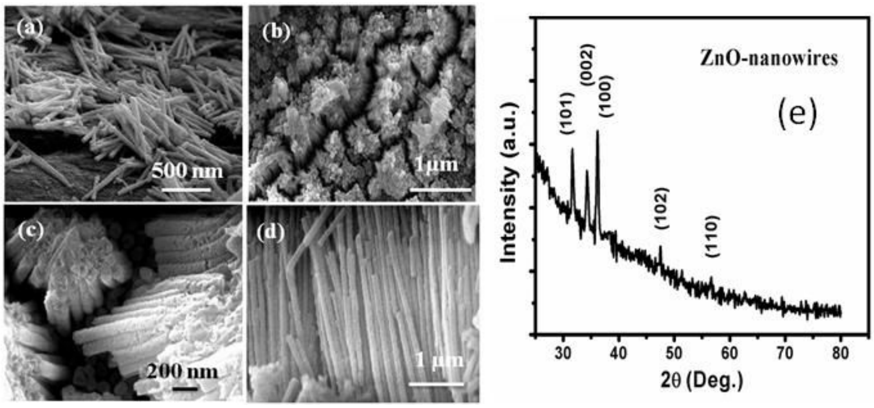


***Fig. 1** – SEM image of ZnO nanowires grown (a) at* 70 °C *in PCM, (b) and (c) at* 80 °C *in AAO membrane and (d) cross-sectional view of ZnO nanowires and (e) X-ray diffraction pattern of ZnO nanowires*

The SEM image of ZnO nanowires/nanorods synthesized at 70 - 90 °C in PCM and AAO membrane is shown in Fig. 1a-d. Fig. 1a shows SEM image of ZnO nanorods grown at 70 °C in PCM template and completely dissolve in dichloromethane. Fig. 1b and c shows the top view of the ZnO nanowires with the AAO template partly dissolve in 1 M NaOH solution. The arrays of nanowires in the image are vertically standing. The cross-sectional image of ZnO nanowires arrays deposited for 30 minute is shown in Fig. 1d. The nanowires are oriented in same direction and some nanowires are broken off due to the mechanical forces during etching of template. The majority of nanowires are approximately 20 μm in length (deposition rate approximately ~ 0.6 μm), which can be modulated by varying the deposition time. In order to obtain high filling of uniform nanowires arrays the membrane is ultrasonicated in an ethanol for 5 min and then dipped in electrolyte for 10 min. The Cd concentration was varied from 0 - 16 at%. X-ray diffraction patterns of ZnO nanowires arrays embedded in AAO template is shown in Fig. 1 e. Strong reflections corresponding to (100), (002) and (101) planes are observed along with the weaker reflections of (102) and (110) planes of wurtzite ZnO, indicating ZnO nanowires are polycrystalline. We have observed from the X-Ray Diffraction patterns that there is no phase separation up to 16 at% of Cd. Fig. 2 shows the ZnCdO nanorods with the average diameter of 100 - 120 nm and average length up 3 - 4 μm. In the primary solution the Cadmium concentration was varied from 0 - 16 at%. No substantial change were observed up to 16 at% in the dominant peak reflections but there is a peak shift towards low-angle side when compared with the pure ZnO. This shift in peak implies that there is a compressive stress in the *c*-axis orientation.

The lattice expansion indicates that the substitution of Zn atom by larger Cd atom. TEM images are shown in Fig. 3. Energy dispersive spectrum confirms the presence of Zn and oxygen in the arrays of nanowires shown in Fig. 3 a. ZnO nanowires grown in AAO/PCM templates are tapered along the

to Contents, page 1

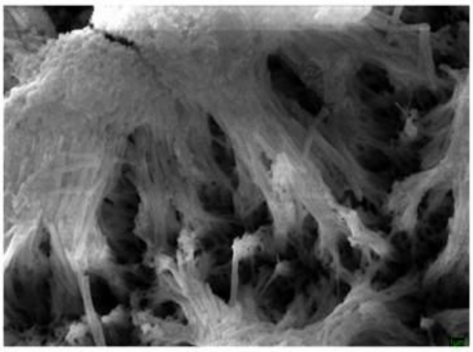


Fig. 2 *– SEM of $Zn_{0.96}Cd_{0.04}O$ Nanorods grown in AAO template at* 80 °C

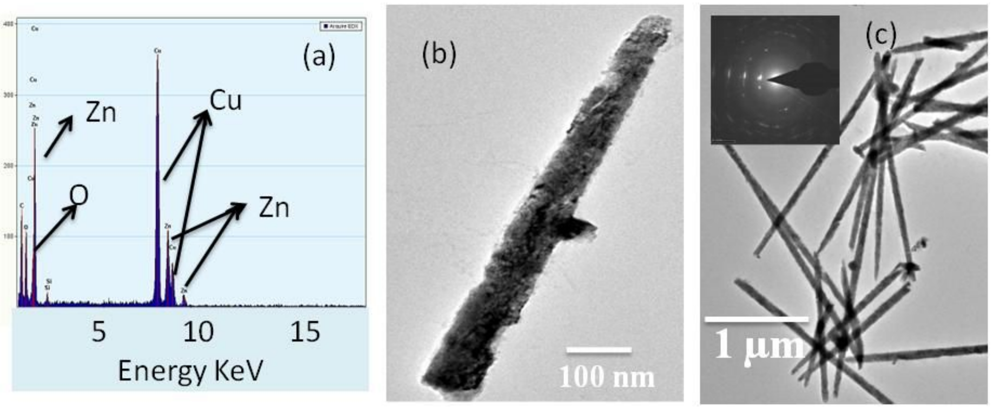


Fig. 3 *– EDS spectrum of ZnO nanowires (a), TEM image of ZnO nanowires at* 70 °C *in PCM (b) and at* 90 °C *(c) (in inset SAED pattern of ZnO nanowires)*

growth direction and shown in Fig. 3 b and c. The nanowires grown by electrochemical method are of polycrystalline nature and selected area electron diffraction pattern is shown in Fig. 3 c.

4. CONCLUSIONS

We have shown that the electrodeposition is an effective technique to synthesize ZnO and ternary ZnCdO nanowires/nanorods in the porous polycarbonate and anodic alumina membrane. The nanowires were grown at a constant potential – 1.0 V and deposition temperature was varied from 70 - 90 °C. XRD analysis showed the ZnCdO nanorods were of pure ZnO wurtzite structures. It was also observed that there was compressive stress in the c-axis orientation. TEM measurements showed the nanowires/nanorods were of good crystalline quality.

REFERENCES

1. M. Lai, D.J. Riley, *Chem. Mater.* **18,** 2233 (2006).
2. L. Huang, J.S. Wind, P.S. Obrien, *Nano Lett.* **3,** 299 (2003).
3. B.B. Lakshmi, K.P. Dorhout, R.C. Matrin, *Chem. Mater.* **9,** 857 (1997).

to Contents, page 1

4. R.C. Matrin, S.L. Van Dyke, H.Z. Cai, B.W. Liang, *J. Am. Chem. Soc.* **112**, 8976 (1990).
5. S. Park, H.J. Lim, W.S. Chung, A.C. Mirkin, *Science* **303**, 348 (2004).
6. A.P. Alivisatos, *Science* **271**, 933 (1996).
7. C.M. Lieber, *Solid State Commun.* **107**, 607 (1998).
8. V.E. Henrich, P.A. Cox, *The Surface Science of Metal Oxides*, (Cambridge University Press: Cambridge: 1994).
9. H. Cao, J.Y. Xu, D.Z. Zhang, S.H. Chang, S.T. Ho, E.W. Seelig, X. Liu, R.P.H. Chang, *Phys. Rev. Lett.* **84**, 5584 (2000).
10. I.D. Makuta, S.K. Poznyak, A.I. Kulak, E.A. Streltsov, *Phys. Status Solidi A* **111**, 193 (1989).
11. C.X. Shan, Z. Liu, Z.Z. Zhang, D.Z. Shen, S.K. Hark, *J. Phys. Chem. B* **110**, 11176 (2006).
12. Y.J. Li, C.Y. Wang, M.Y. Lu, K.M. Li, L.J. Chen, *Cryst. Growth Des.* **8**, 2598 (2008).
13. F. Bertram, S. Giemsch, D. Forster, J. Christen, R. Kling, C. Kirchner, A. Waag, *Appl. Phys. Lett.* **88**, 061915 (2006).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.151-154

PACS numbers: 61.43Fs, 64.70.P, 61.40D, 81.70.Pg, 61.43Dq, 65.60

KINETICS STUDY OF $(Se_{80}Te_{20})_{100-x}Cd_x$ GLASSY ALLOY BY DIFFERENTIAL THERMAL ANALYSIS USING NON-ISOTHERMAL TECHNIQUE

Mainika, N. Thakur

Department of Physics, H.P. University, Summer Hill, Shimla-5
Email: mainika_phy@rediffmail.com

The kinetics of crystallization in $(Se_{80}Te_{20})_{100-x}Cd_x$ (x = 0, 2, 4 and 6) alloys at different heating rates have been studied by Differential Thermal Analysis in non-isothermal condition. A comparison of various quantitative methods to assess the level of stability of the glassy material in the above mentioned system is presented. All these methods are based on the characteristics temperature obtained by heating of the samples, such as glass transition temperature (T_g), temperature of crystallization (T_c), and the melting temperature (T_m). From the dependence of glass transition temperature on heating rate, the activation energy (E_g) has been calculated on the basis of the Kissinger and Moynihan models.

***Keywords**: CHALCOGENIDE GLASSES, ACTIVATION ENERGY, T_g, DTA, Se-Te-Cd.*

(Received 04 February 2011)

1. INTRODUCTION

Thermal analysis has been extensively used for studying the kinetics of chemical reaction [1-2] and crystallization of glasses [3-8]. The concepts of kinematical studies are always connected with the activation energy. In calorimetric measurements, two basic methods can be used: isothermal and non-isothermal [9-13]. In the isothermal method, the sample is brought quickly to a temperature above the glass transition temperature and the heat evolved during the crystallization process at a constant temperature is recorded as a function of time while in the non-isothermal method, the sample is heated at a fixed rate (β) and the heat evolved is recorded as a function of temperature or time [14]. The main factor, which leads to the stability of amorphous phase of a compound near room temperature, is its glass transition temperature, if the melting temperature of the compound is high, and the glass transition temperature, T_g, of it is expected to be considerably higher than room temperature [15]. The activation energy plays a dominant role in deciding the utility of the material for the specific purpose. One of the most important aspects of the study of glasses is the composition dependence of properties. Among amorphous chalcogenide alloys, selenium based melt are characterized by high viscosity [16]. This feature favors the glass formation in bulk form by air-quenching or water-quenching as well as in evaporated thin film forms. Since tellurium based melts with the same elements generally have low viscosity, a high cooling rate is required to prevent nucleation and growth during quenching and to obtain bulk glasses. Glassy alloys of the Se-Te system have become materials of considerable commercial, scientific and technological importance as they

to Contents, page 1

have greater hardness, higher crystallization temperature, higher photosensitivity and smaller ageing effects than pure Se [17]. The addition of a third element (Cd) expands the glass forming area and also creates compositional and configurational disorder in the system. Present paper is concentrated on kinetic studies of $(Se_{80}Te_{20})_{100-x}Cd_x$ (x = 0, 2, 4, 6) glassy alloy under non-isothermal technique by using DTA.

2. EXPERIMENTAL PROCEDURE

For preparation of Se-Te-Cd glasses, high purity (5N) elements in appropriate atomic percentages were weighted into the quartz ampoules and sealed off in a vacuum of 10^{-5} torr and then heated in furnace at around 1000 °C for 12h. The ampoules were rotated frequently to ensure homogenization. The ampoules were then rapidly quenched in ice-water to obtain the glasses and samples were removed by breaking the quartz ampoule. SHIMADZU DTG-60, simultaneous TG/DTA module is used to measure the caloric manifestation of the phase transformation. The TG/DTA scans are taken at four heating rate (10, 15, 20, 25°C/min) for the four different composition in the micro alumina pans under dry nitrogen gas. The masses of the samples varied between10 and 30 mg. The amorphous nature of the resulting glassy alloys was verified by X-ray diffraction as no prominent peak was observed.

3. RESULT AND DISCUSSION

The glass transition temperature represents the strength or rigidity of the glassy structure of the alloy. In present work glass transition region has been studied in terms of variation of glass transition temperature with the heating rate. Three approaches are used to analyze the dependence of T_g on the heating rate (β). The first is the empirical relationship, which has originally been suggested by Lasocka [18] and has the form,

$$T_g = A + B \ln\beta \quad (1)$$

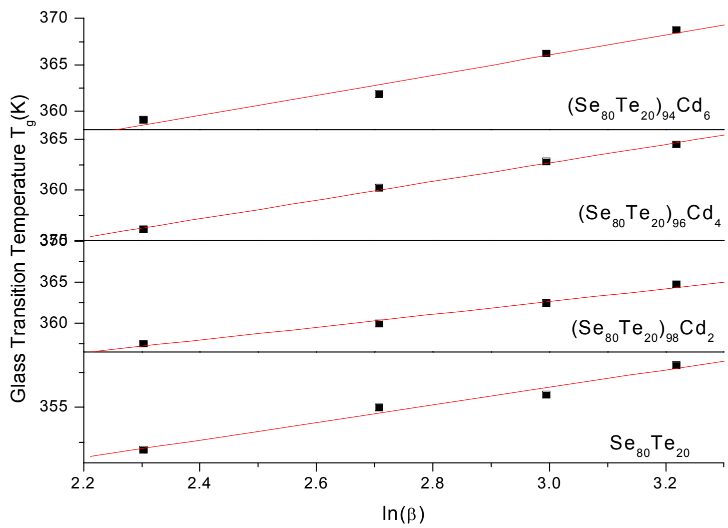


Fig. 1 *– Plots of T_g versus lnβ*

to Contents, page 1

where A and B are constants for a given glass compositions and calculated values are given in Table 1. Plots of T_g versus logβ indicates the validity of Eq. (1) for our compositions (Fig. 1).

The activation energy (E_g) of the glass transition process is one of the most important parameters for understanding the thermal relaxations that occur in glassy networks during the glass transition. The activation energy for glass transition (E_g), depends on T_g as a function of the heating rate (β) have been evaluated using Kissinger's equation [19-21] and is given by the following expression,

$$\ln(\beta/T_g^2) = -E_g/RT_g + \text{constant} \quad (2)$$

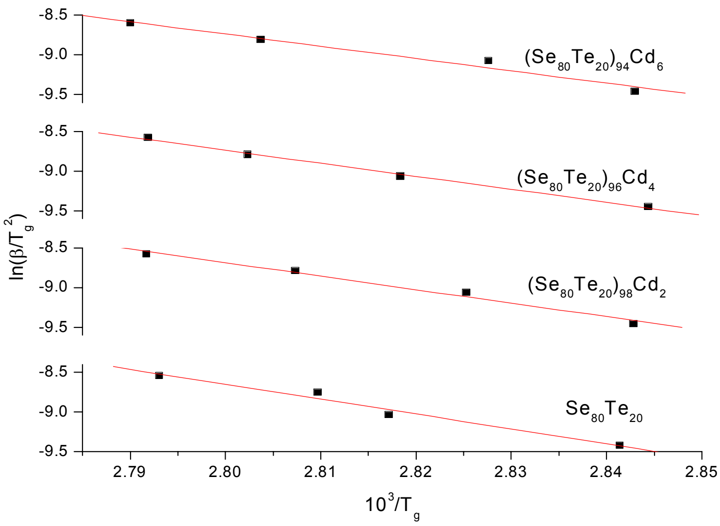


Fig. 2 *– Shows the plots of* $ln(\beta/T_{g2})$ *versus* $10^3/T_g$

The activation energy of the glass transition (E_g) also has been evaluated using the Moynihan [22] relation derived, based on the concept of thermal relaxation.

The plot of $\ln\beta$ versus $10^3/T_g$ yields a straight line, the slope of which gives the activation energy of glass transition (Fig. 3). Table 1 lists the values of E_g obtained using the Kissinger and Moynihan models, which are in good agreement with each other with the difference within experimental error. The glass transition activation energy is the amount of energy that is absorbed by a group of atoms in the glassy region so that a jump from one metastable state to another state is possible [23]. Accordingly, the atoms in a glass having minimum activation energy have a higher probability to jump the metastable (or local minimum) state of lower internal energy and, hence, are the most stable.

Table 1 *–* E_g *obtained from the Kissinger and Moynihan Models (kJ/mol)*

Composition	Kissinger Model	Moynihan Model	A(K)	B(K)
$x = 0$	155.88	161.86	337.48	6.32
$x = 2$	131.09	137.09	339.25	7.80
$x = 4$	110.64	116.62	335.06	9.21
$x = 6$	92.67	98.72	333.51	10.83

to Contents, page 1

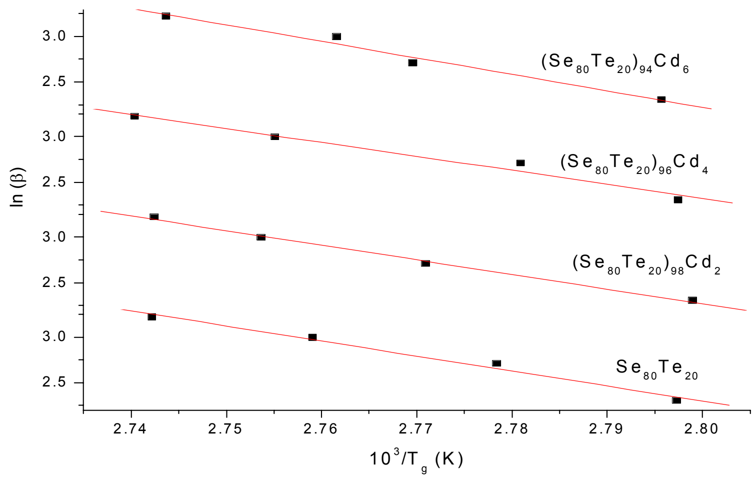


Fig. 3 *– The plot of lnβ versus $10^3/T_g$*

4. CONCLUSION

The values of activation energies for glass transition were found to decrease with increase in Cd content in Se-Te glassy alloy. The values of activation energies, using two different methods, are in good agreement with each other. So it can be concluded that any of these two methods can be used to calculate glass transition activation energy.

REFERENCES

1. G.O. Piloyan, I.D. Ryabchikov, O.S. Novikova, *Nature* **212,** 1229 (1966).
2. S. Tanabe and R. Otsuka, *Netsu Sokutei. (Calor. Therm. Anal.)* **4**, 139 (1977).
3. R.L. Thakur, S. Thiagrajan, *Bull. Cent. Glass Ceram. Res. Inst.* **13**, 33 (1966).
4. R.L. Thakur, S. Thiagrajan, *Bull. Cent. Glass Ceram. Res. Inst.* **15**, 67 (1968).
5. K. Matusita, S. Sakka, *Phys. Chem. Glasses* **20**, 81 (1979).
6. K. Matusita, S. Sakka and Y. Matsui, *J. Mater. Sci.* **10,** 961 (1975).
7. A. Pratap, A. Prasad, S.R. Joshi, N.S. Saxena, M.P. Saksena, K. Amiya, *Mater. Sci. Forum* **179-181,** 851 (1995).
8. A. Pratap, K.G. Raval, A. Gupta, S.K. Kulkarni, *Bull. Mater. Sci.* **23,** 185 (2000).
9. J. Vazquez,P.L. Lopez-Alemany, P. Villares, R. Jimenez-Garay, *J. Phys. Chem. Solids* **61,** 493 (2000).
10. M.A. Abdel-Rahim, *Physica B* **239,** 238 (1997).
11. A.H. Moharram, M.M. Hafiz, A.A. Abu-Sehly, *Appl. Surf. Sci.* **137,** 150 (1999).
12. A.H. Moharram, M.S. Rasheedy, *Phys. Status Solidi a* **169,** 33 (1998).
13. M. Abu El-Oyoun, *J. Phys. D: Appl. Phys.* **33,** 2211 (2000).
14. M. Saxena, P.K. Bhatnagar, *Bull. Mater. Sci.* **26,** 547 (2003).
15. N. El-Kabany, *Chalcogenide Lett.* **7**, 351 (2010).
16. M. Ilyas, M. Zulfequar and M. Husain, *Opt. Mater.* **13,** 397 (2000).
17. S.O. Kasap, T. Wagner, V. Aiyah, O. Krylouk, A. Bekirov, L. Tichy, *J. Mater. Sci.* **34,** 3779 (1999).
18. Maria Lasocka, *Mater. Sci. Eng.* **23,** 173 (1976).
19. H.E. Kissinger, *Anal. Chem.* **29,** 1702 (1957).
20. H.S. Chen, *J. Non-Cryst. Solids* **27,** 257 (1978).
21. N.A. Macmillan, *J. Phys. Chem.* **42**, 3497 (1965).
22. C.T. Moynihan, A.J. Easteal, J. Wilder, J. Tucker, *J. Phys. Chem.* **78,** 2673 (1974).
23. J. Fusong, M. Okuda, *Jpn. J. Appl. Phys.* **30,** 97 (1991).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.155-161

PACS numbers: 61.46.Bc, 68.37.Og, 78.30.Ly 78.67.Hc, 81.07.Bc

A STUDY OF THE EVOLUTION OF THE SILICON NANOCRYSTALLITES IN THE AMORPHOUS SILICON CARBIDE UNDER ARGON DILUTION OF THE SOURCE GASES

A. Kole, P. Chaudhuri

Energy Research Unit, Indian Association for the Cultivation of Science,
Jadavpur, Kolkata – 700032
E mail: erak@iacs.res.in

Structural evolution of the hydrogenated amorphous silicon carbide (a-SiC:H) films deposited by rf-PECVD from a mixture of SiH_4 and CH_4 diluted in Ar shows that a smooth transition from amorphous to nanocrystalline phase occurs in the material by increasing the Ar dilution. The optical band gap (E_g) decreases from 1.99 eV to 1.91 eV and the H-content (C_H) decreases from 14.32 at% to 5.29 at% by increasing the dilution from 94 % to 98 %. at 98 % Ar dilution, the material contains irregular shape Si nanocrystallites with sizes over 10 nm. Increasing the Ar dilution further to 98.4 % leads to a reduction of the size of the Si nanocrystals to regular shape Si quantum dots of size about 5 nm. The quantum confinement effect is apparent from the increase in the E_g value to 2.6 eV at 98.4 % Ar dilution. Formation of Si quantum dots may be explained by the etching of the nanocrystallites of Si by the energetic ion bombardment from the plasma.

Keywords: *SILICON CARBIDE, RF- PECVD, AR DILUTION, OPTICAL BAND GAP, SI QUANTUM DOTS, QUANTUM CONFINEMENT.*

(Received 04 February 2011)

1. INTRODUCTION

Study of the formation of silicon quantum dots (Si q-dots) which are silicon nanocrystallites having sizes of the order of Bohr atomic radius (~5nm) embedded in high band gap amorphous oxides, nitrides or carbides of silicon is highly interesting from the point of view of understanding the basic physics of the interaction of the light quanta with such quantum size matter. Tuning of the band gap in these materials by controlling the size of the Si q-dots have possible technical application in photoluminescent devices and in for third generation solar cells.[1, 2] Most of the works on Si q-dots in a-SiC matrix used hydrogen dilution of the source gases containing Si and C.[3] In this paper we are reporting our studies on the gradual evolution of the nanocrystalline Si into Si quantum dots in a-SiC:H films by the control of the argon dilution of the source gases.

2. EXPERIMENTAL

The a-SiC:H films were deposited by the conventional rf PECVD technique from a mixture of SiH_4, CH_4, and Ar. Total flow of the process gases was maintained at 100 sccm. SiH_4, and CH_4 were flown at 1:1 ratio and the Ar dilution given by $[Ar]\cdot 100\ \%/([Ar] + [SiH_4] + [CH_4])$ was varied from 94 %

to Contents, page 1

to 98.4 %. The preparation conditions of the samples are given in Table 1. The crystalline structure of the nanophase films was investigated using a Seifert XDAL 3000 X-ray diffractometer, operating in the grazing incident geometry (incident angle of 2°).The incident X-ray wave length was 1.5418 Å (Cu Kα line) at 35 kV and 30 mA. The nanostructure of the films was studied by HRTEM (JEOL 2010). The hydrogen and carbon bondings with silicon was studied by FTIR absorption spectroscopy in the frequency range between 400 and 4000 cm^{-1}. The optical absorption and the band gap of the films were measured by UV-vis spectrophotometer (HITACHI U4100).

Table 1 – *Preparation conditions of the samples*

Sample No.	SiH_4 (sccm)	CH_4 (sccm)	Ar (sccm)	Temp (°C)	Press (Torr)	Power density (mW/cm^2)
#QD1	3	3	94	200	0.2	400
#QD2	2	2	96	200	0.2	400
#QD3	1.5	1.5	97	200	0.2	400
#QD4	1	1	98	200	0.2	400
#QD5	0.8	0.8	98.4	200	0.2	400

3. RESULTS

3.1 XRD

No diffraction peak related to crystalline silicon is observed up to a dilution level of 96 % (Curves 1 and 2, Fig. 1). But at 97 % dilution of Ar small, broad but clearly resolved diffraction peaks at $2\theta = 28.4°$, 47.3°, and 56.1° appear corresponding to the (111), (220), and (311) crystal planes of silicon respectively (Curve 3, Fig. 1).[4, 5] With further increase of the dilution level to 98 %, the intensities of all three diffraction peaks are increased while the full widths at half maximum (FWHM) of all these diffraction peaks become smaller [Curve 4, Fig. 1]. An additional peak at 69.2° corresponding to (400) plane of Silicon also appears at this dilution level. The peaks however, becomes broad when the Ar dilution level is further increased to 98.4 % (Curve 5, Fig. 1).

3.2 Optical band gap

The optical absorption of the films were measured by UV-Vis spectroscopy to study the optical band gap of the materials. The band gap was obtained from the commonly used Tauc's formula (equation 1),

$$(\alpha h\nu)^{0.5} = B(h\nu - E_g), \tag{1}$$

where α is the optical absorption coefficient, B is the joint optical density of states, and $h\nu$ is the incident photon energy. [6,7] Fig. 2 shows the Tauc's plot for different diluted films to determine the optical band gap E_g which was obtained from the plot of $(\alpha h\nu)^{0.5}$ versus $h\nu$ by extrapolating the linear portion of the curve to intercept the energy axis(at $\alpha = 0$). It is observed that with increasing Ar dilution from 94 % to 98 % there is a continuous decrease of the band gap from 1.99 eV to 1.91 eV. Increasing the dilution further to 98.4% there is a sharp increase in the band gap to 2.6 eV.

to Contents, page 1

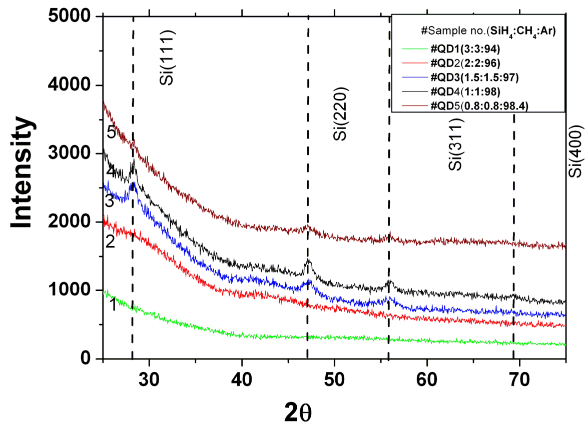


***Fig. 1** – XRD pattern of the films deposited under different dilution levels of Ar*

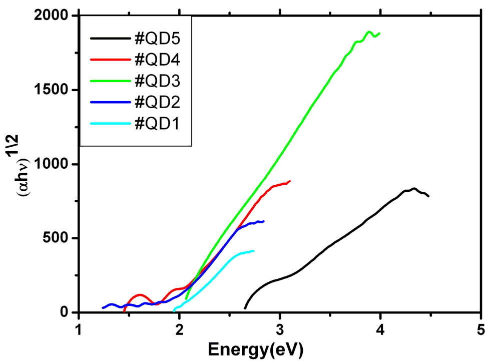


***Fig. 2** – Tauc's plot of the a-SiC:H samples deposited at different ar dilution*

3.3 FTIR

The nature of the Si-C, Si-H and C-H bonds within the samples deposited at different dilution level of Ar studied by FTIR absorption spectroscopy are shown in Fig. 3. These spectra have been corrected for the substrate absorption and normalized by the film thickness. The main absorption peaks appearing in the spectra are located at (1) ~ 650cm^{-1} , (2) ~ 780 cm^{-1}, (3) ~ 1000 cm^{-1}, (4) 1900 - 2100 cm^{-1} and (5) 2800 - 3100 cm^{-1}.These peaks are attributed to (1) the wagging or rocking mode of Si-H$_n$ for $n = 1$ - 3, [8], (2) the stretching mode of Si-C, [9], (3) the wagging or rocking mode of C-H$_n$ [10, 11], (4) the stretching mode of Si-H [12] and (5) stretching modes of C-H$_n$ (sp^3) or C-H$_n$ (sp^2) for $n = 1$ - 3 [13] respectively. With the increase in Ar dilution level the intensities of the peaks at ~ 650 cm^{-1} and ~ 2090 cm^{-1} related to Si-H bond decrease. This phenomenon may be associated with the nanocrystalline Si formation in the amorphous matrix. [14] It is also observed that the peak at ~ 1000 cm^{-1} and the absorption band between 2800 -3100 cm^{-1} (both related to C-H$_n$ mode) decrease with the increase in Ar dilution level. This observation together with the decrease

to Contents, page 1

of the Si-C stretching mode at 780 cm^{-1} indicate that the a-SiC:H films are getting Si rich with the increase in Ar dilution level. Hydrogen content (C_H) within the films estimated by de-convoluting the absorbance of Si-H wagging or rocking modes are shown in Fig. 5. It is clear from the graph that with the increasing dilution level the Hydrogen content within the films decreases rapidly which is an important indication of the formation of nano crystallites within the films.

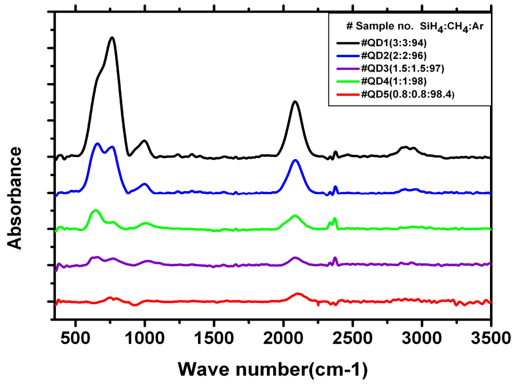


Fig. 3 *– FTIR absorption spectra of the films grown at different Ar dilution*

3.4 HRTEM

At 94% Ar dilution amorphous nature of a-SiC:H films is apparent from the featureless micrograph (See Fig. 4a). Corresponding SAED (Fig. 4b) shows diffused hallow pattern. With the increase in Ar dilution level diffused rings appear in the SAED indicating formation of the nanocrystallites. The micrograph and the SAED for the 98 % diluted sample are shown in Fig. 4c, d. The micrograph shows randomly oriented nanocrystalline Si of variable sizes. The fringe pattern corresponding to (111) plane of silicon is clearly discernible in the micrograph (Fig. 4c inset). The SAED (Fig. 4d) also shows a sharp ring corresponding to this plane. At 98.4 % Ar dilution the micrograph consists of uniformly distributed nanocrystals of size ~ 5 nm Fig. 4e). A distinct feature is observed in the SAED of this film showing a diffused ring with some spots on it. The distinct spots coincide with the Laue spots for the planes (220) and (311) of Si. This unique feature may have appeared in the diffraction pattern when the size of the nanoparticle is reduced to the size of the order of Bohr radius producing the Si quantum dots in the dielectric matrix of SiC.

4. DISCUSSIONS

Formation of Si quantum dots in a host matrix may be achieved along two routes; firstly by the “bottom up” method where the gradual build up of the nanocrystals of the size of “Bohr radius” is achieved by assembling of the Si atoms in a regular arrangement and secondly, by “top down” method where the larger size nanocrystals are etched out to reach the size when the quantum confinement effect will be observed. The XRD data shows a gradual increase of the peaks corresponding to the various planes of

to Contents, page 1

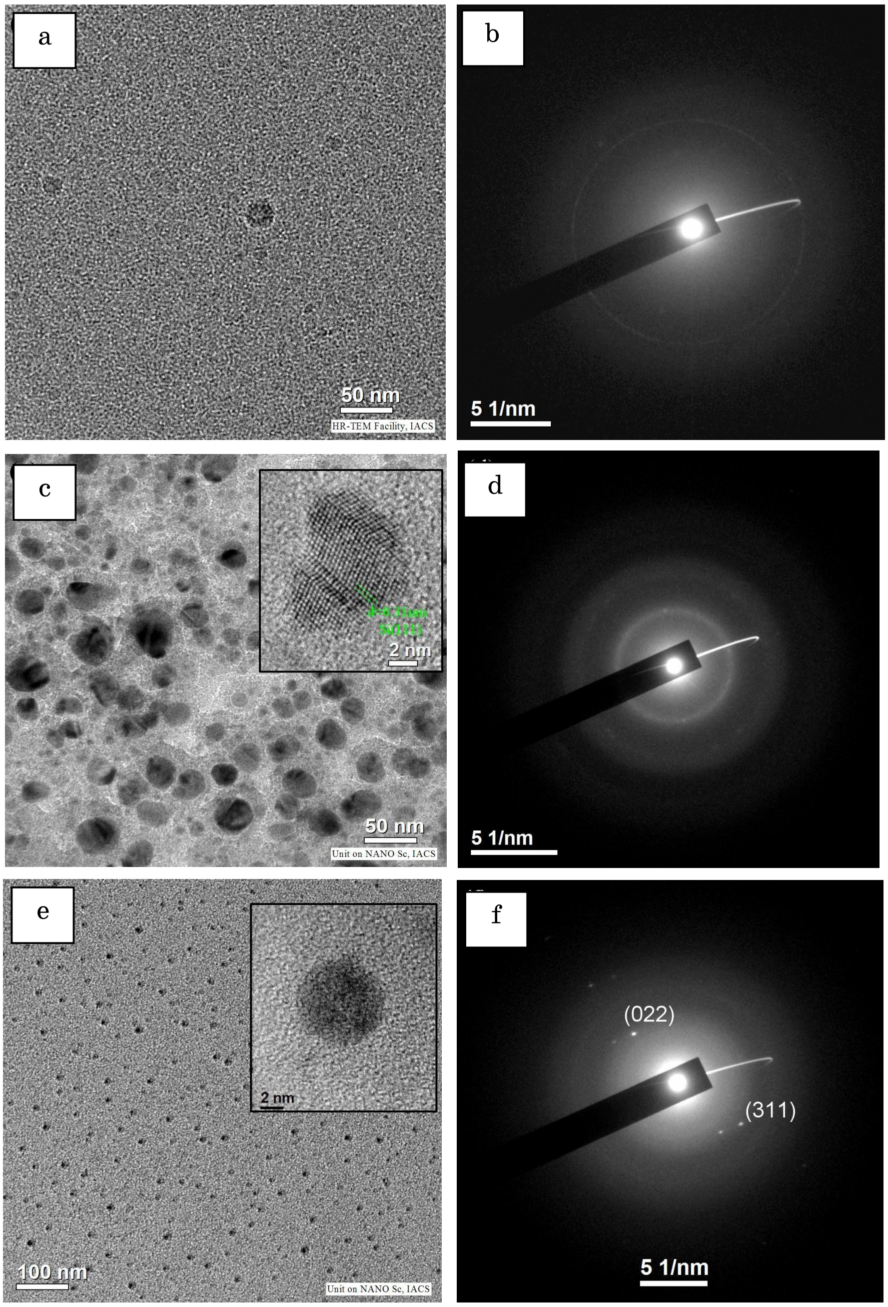


Fig. 4 *– HRTEM images of the samples deposited at different Ar dilution. (a), (c), (e) shows the bright field micrograph of the samples deposited at* 94 %, 98 % *and* 98.4 % *respectively. (b), (d) and (f) shows the corresponding selected area electron diffraction (SAED) pattern of the samples for* 94 %, 98 % *and* 98.4 % *Ar dilution levels respectively*

to Contents, page 1

crystalline silicon which indicates increase of the size of the Si nanocrystallites in the a-SiC:H matrix with increase of of Ar dilution from 94 % to 98 %. Increase in the size of the Si nanocrystallites within a-Si has been observed with increasing the dilution of the source gas (SiH_4) with hydrogen has been reported [13]. In this study we observe a broadening of the XRD peaks at the highest dilution level of 98.4 % (Curve 5, Fig. 1). The TEM micrograph of the 98 % Ar diluted sample shows irregular shape large size (> 10 nm) Si nanocrystallites are observed. Corresponding SAED shows rings for Si(111) plane. On increasing the Ar dilution to 98.4 % decrease of the size of the Si nanocrystallites is evident from the inset of the micrograph (Fig. 4e). Corresponding SAED consists of diffused ring with some Laue spots.

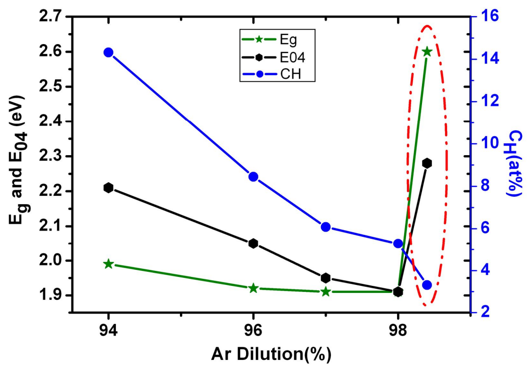


***Fig. 5** – Changes in C_H, E_{04} and E_g with Ar dilution*

Argon takes leading role in the dissociation of SiH_4 and CH_4 in the plasma as also influencing the surface reactions. While the Ar^+ ions have higher reaction rate with CH_4, the neutral metstable Ar^* is chiefly responsible for the dissociation of SiH_4. Thus presence of the relative amounts Ar^* and Ar^+ in the plasma determines the dissociation of SiH_4 and CH_4. Moreover, bombardment of the growth surface by Ar^+ and Ar^* influences the surface reactions which is important for the evolution of the deposited from amorphous to nanocrystalline phase. Diffused rings containing Laue spots have been observed in the case of Si quantum dots formed within porous silicon [15].

A strong evidence of the Si quantum dot formation in the a-SiC:H matrix is obtained by studying the variation of the E_g and E_{04} with the Ar dilution (Fig. 5). Increase of the size of the Si nanocrystallites causes a decrease of E_g from 1.99 eV to 1.91 eV and E_{04} from 2.21 eV to 1.91 eV by increasing the Ar dilution from 94 % to 98 %. With further increase in the Ar dilution to 98.4% both E_g and E_{04} sharply rises to the values of 2.6 eV and 2.28 eV respectively. Such increase in the band gap can not be explained by the alloying of Si with C or H because both C and H bonding decrease with increasing Ar dilution. Bonded H-content C_H in the material with change in Ar dilution plotted in Fig. 5 show a continuous decrease from 14.32 at% to 3.31 at% with increase in Ar dilution from 94 % to 98.4 %. Reducing the size of the Si crystallites to the size of the Bohr atomic radius (~ 5 nm) has been found to increase the band gap due to quantum confinement effect.

to Contents, page 1

With increasing Ar dilution the growth of amorphous to micro/nano-crystalline silicon occurs through the more and more bombardment of the growth surface by the metastable Ar^* atoms and Ar^+ ions from the plasma. Energy transferred to the surface increases the mobility of the surface adatoms helping in the formation of the nanocrystallites [16]ю At higher Ar dilution the momentum transfer to the surface causes etching from the surface. The nanocrystallites are also etched and their size reduced to form silicon quantum dots.

5. CONCLUSION

We have observed that deposition of quantum dots of uniform and regular size in a-SiC:H matrix occurs through the etching out of the initially formed larger size and irregularly shaped nanocrystallites of silicon. Bombardment of the growth surface by the ions from the plasma has been proposed to play a major role in the etching process.

REFERENCES

1. S.M. Kang, S.G. Yoon, S.W. Kim, and D.H. Yoon, *J. Nanosci. Nanotechno.* **8** 3857 (2008).
2. G. Conibeer, M. Green, R. Corkish, Y. Cho, E.-C. Cho, Chu-W. Jiang, T. Fangsuwan-narak, E. Pink, Y. Huang, T. Puzzer, T. Trupke, B. Richards, A. Shalav, K.-L. Lin, *Thin Solid Films* **511,** 654 (2006).
3. Y. Kurokawa, S. Tomita, S. Miyajama, A. Yamada, M. Konagai, *Jap. J. Appl. Phys.* **46** L833 (2007).
4. Q.J. Cheng, S. Xu, S.Y. Huang, K. Ostrikov, *Cryst. Growth Des.* **9** 2863 (2009).
5. Y. Hotta, H. Toyoda, H. Sugai, *Thin Solid Films* **515** 4983 (2007).
6. Q.J. Cheng, S. Xu, K. Ostrikov, *J. Phys. Chem. C* **113** 14759 (2009).
7. Q.J. Cheng, S. Xu, K. Ostrikov, *J. Mater. Chem.* **19** 5134 (2009).
8. M.H. Brodsky, M. Cardona, J.J. Cumo, *Phys. Rev. B* **16,** 3556 (1977).
9. Y. Catherine and G. Turban, *Thin Solid Films* **60** 193 (1979).
10. H. Wieder, M. Cardona, and C.R. Guarnieri, *phys. status solidi B* **92,** 99 (1979).
11. W. Yu, X. Wang, W. Lu, S. Wang, Y. Biana, G. Fu, *Physica B* **405** 1624 (2010).
12. A. Bhaduri, A. Kole, P. Chaudhuri, *phys. status solidi C* **7,** 774 (2010).
13. Q. Cheng, E. Tam, S. Xu, K. Ostrikov, *Nanoscale* **2,** 594 (2010).
14. A. Chowdhury, S. Ray, *phys. status solidi C* **7,** 628 (2010).
15. Z. Yamani, O. Gurdal, A. Alaql, M.H. Nayfeh, *J. Appl. Phys.* **85,** 8050 (1999).
16. U.K. Das, P. Chaudhuri, S.T. Kshirsagar, *J. Appl. Phys.* **80,** 5389 (1996).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.162-169

PACS numbers: 64.70.kg, 65.40.gh

EFFECTS OF INTERFACIAL CHARGES ON DOPED AND UNDOPED HfO_x STACK LAYER WITH TIN METAL GATE ELECTRODE FOR NANO-SCALED CMOS GENERATION

***S. Chatterjee*[1], *Y. Kuo*[2]**

[1] ECE Department, Techno India,
EM-4/1, Sector-V, Salt Lake, Kolkata, 700091, India
E-mail: somenath@gmail.com

[2] Thin Film Nano & Microelectronics Research Laboratory,
Texas A&M University,
College Station, TX 77843-3122, USA

A comparison of the interfacial charges present in the high-k stacked gate dielectrics for Zr-doped HfO_x and undoped HfO_x samples with titanium nitride (TiN) metal gate electrode is reported here. The metal gate work function value (4.31 eV) for TiN gate electrode was extracted from the TiN/SiO_2 /p-Si capacitor. The calculated charge densities in both doped and undoped films are of the order of 10^{12} cm^{-2}. The interfacial charge present in the high-k/SiO_2 interface is negative for ALD deposited pure HfO_2 samples; where as the charges are positive for RF-sputter deposited pure HfO_2 and Zr-doped HfO_x samples. The existence of positive interface charges may be due to the fabrication process.

Keywords: *DOPED HIGH-K GATE DIELECTRICS, NANOELECTRONICS, MOSFET, WORK FUNCTION, OXIDE-DEFECTS.*

(*Received 04 February 2011*)

1. INTRODUCTION

Fabrication of smaller and faster metal-oxide-semiconductor (MOS) field effect transistors and superior memory devices dictate the usage of new materials and chemical processes to make nano-electronics a reality. The high gate leakage current, doping penetration and gate depletion effect will limit the use of SiO_2 (as gate dielectrics) and poly-Si (as gate electrodes) for sub-65 nm technology node. Alternative dielectric materials such as Si_3N_4, HfO_2, [1] ZrO_2, [2] their silicates and transition metal doped high-k dielectrics [3, 4] have been suggested as candidate materials. The amorphous-to-polycrystalline phase transition temperature of the film can be increased by adding a third element into the oxide, e.g. Zr doped in the HfO_x, because of the polycrystalline high-k film degraded the device reliability due to the uneven distribution of grains in the channel region [3]. Furthermore, the doped element can control the fast diffusion of oxygen vacancies in high-k films, which is mainly responsible for the formation of uncontrolled interfacial layer thickness, lower breakdown field and higher leakage current density [3-5]. There is no literature so far concerning the gate dielectrics with Zr^{4+} doped HfO_x. It is well known that Zr and Hf are both 4-valence elements, so Zr doped HfO_x would not exhibit any increase in oxygen voids in the film.

to Contents, page 1

In contrast to the high-k dielectric selection that is nearing consensus, the searching of metal gate electrode for CMOS is in its infancy. One of the requirements for the integration issues of a new metal gate electrode is the proper set of work function values. In the PMOS (NMOS) transistor, heavily p-type (n-type) doped poly-Si is used as gate electrode and the work function is about 5.2 eV (4.1 eV). So, the substituting metals or metal compounds should have the work function, i.e. the gate fermi level for PMOS (NMOS) devices is 0.2 eV above (below) the band edge E_v (E_c), in order to reduce the transistor's threshold voltage [6]. The metal gate work functions depend on bulk and surface material properties, crystalline orientation and the permittivity of the dielectric interfacing with the metal. The work function of a metal at a dielectric interface is different from its value in vacuum. This may be explained either by metal induced gap states (MIGS), as the interface provides the possibilities such as metal-insulator transition or by the formation of a dipole layer at the metal-dielectric interface [7]. Again the defects/charges present in the high-k gate dielectrics stack are different than the defects present in the conventional SiO_2 gate dielectrics in SiO_2/Si system. Moreover, the present of charges will lead to large shift in transistor threshold voltage. Columbic scattering from excess charge in the high−k film will most certainly cost degradation in channel carrier mobility to unacceptable levels. Therefore, an investigation of the impact of interfacial charges on the metal gate's work function with doped high-k dielectric stack is technically important and timely.

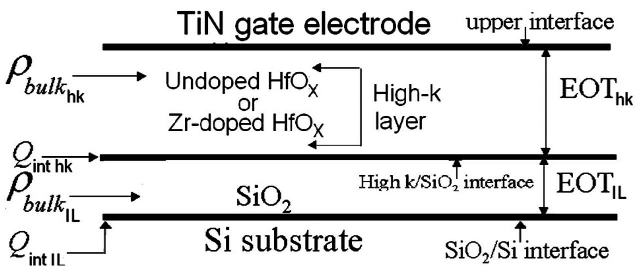


***Fig. 1** – The high-k gate dielectric consists of the stacked structure and the charges are situated at 1) upper interface, i.e. between the gate electrode and high-k gate dielectrics, 2) bulk high-k layer, 3) interface between the high-k and SiO_2-rich dielectrics, 4) bulk SiO_2 dielectrics layer, and 5) between the SiO_2 and substrate interface. The EOT_{hk} and EOT_{IL} are the equivalent oxide thicknesses for High-k and SiO_2 layer, respectively*

Fig. 1 shows a schematic diagram of MOS structure with stacked gate dielectrics (detail descriptions are given in the figure caption). $\rho_{bulk\ hk}$ and $\rho_{bulk\ IL}$ are the bulk oxide charges per unit volume in the bulk of high-k and SiO_2 layer, respectively. Similarly, $Q_{int\ hk}$ and $Q_{int\ IL}$ are the fixed sheet charge (charge/area) at the high-k and SiO_2 interface and SiO_2 and Si (substrate) interface. The location of different bulk and interfacial charges, e.g. $\rho_{bulk\ hk}$, $\rho_{bulk\ IL}$, $Q_{int\ hk}$, and $Q_{int\ IL}$, are shown in the Fig. 1.

In this paper, authors have reported the work function for TiN metal nitride gate electrodes with the conventional SiO_2 gate dielectrics, comparison of the different interface charges located in the HfO_2 and Zr-doped HfO_x gate dielectric stacks with TiN gate electrodes.

to Contents, page 1

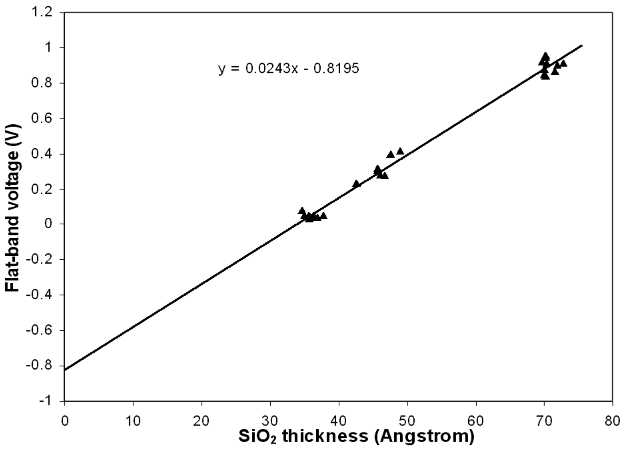


Fig. 2 *– Flat band voltage versus EOT for extraction of metal gate work function using TiN/SiO_2/p-Si MOS capacitor structure.*

2. EXPERIMENTAL

Extraction of the proper work function requires minimization of charges in the oxide, which could vary from wafer to wafer when oxidized to different thicknesses. To avoid such issue, wafer with thick oxide can be partially wet etched in a solution of buffered HF in steps across the wafer in order to achieve the multiple oxide thicknesses (for example 2, 4, 6 nm). The variable thermally grown SiO_2 thicknesses on p-type Si substrate (terraced oxide) and the fixed HfO_2 film thickness using atomic layer deposition (ALD) on the terraced oxide were provided by Sematech [8]. The Zr doped HfO_2 films were deposited by magnetron reactive RF co-sputtering technique using Zr (24 W) and Hf (60 W) metal targets in Ar/O_2 ambient for 20 sec on the terraced oxide substrate. Process parameters such as the reaction time, gas flow rate, and sputtering power were investigated under various conditions to stabilize and optimize process conditions. The post deposition annealing (PDA) was performed for every sample with 700°C in N_2 ambient for 10s using a high temperature substrate heater. TiN metal nitride gates were deposited on the dielectrics by magnetron reactive RF sputtering system in a mixture of Ar and N_2 (50:1) at 5 mTorr for 25 minutes. Post metal annealing (PMA) was done at 425 °C in forming gas (pressure 10 Torr) for 10 mins. For comparison, the undoped HfO_2 films were deposited by magnetron reactive RF sputtering technique on the terraced oxide substrate (20 sec, O_2/Ar mixtures, pressure 5 mTorr). The gate electrode area was defined by photolithography and etched with a mixture of NH_4OH, H_2O_2 and H_2O (5:1:1). For a good ohmic contact of the MOS capacitor, aluminum (Al) film was deposited (using DC Sputter technique) on the backside of the Si after removal of native oxide with HF solution. The capacitance-voltage (C-V) of the MOS capacitor was measured at high frequency (100 kHz) using Agilent 4284A precision LCR meter. The flat-band voltage (V_{FB}) and the equivalent oxide thickness (EOT) of the MOS capacitor were calculated by fitting the high-frequency C-V measurements using a C-V simulation program, developed by NCSU [9].

to Contents, page 1

3. RESULTS AND DISCUSSION

The metal-semiconductor work function difference, ϕ_{ms}, can be estimated using the following equation for TiN metal gate electrodes with different SiO_2 gate dielectrics thicknesses,

$$V_{FB} = \varphi_{ms} - \frac{Q}{C_i} = \varphi_{ms} - \frac{Q\, d_{ox}}{\varepsilon_0 \varepsilon_{SiO_2}} \tag{1}$$

Here, the Q represents the equivalent oxide charge per unit area present in the dielectrics. C_i is the oxide capacitance/area, ε_{SiO2} is the dielectric constant of SiO_2 (~ 3.9), ε_0 is the permittivity of the free space ($8.85 \cdot 10^{-12}$ F/m) and d_{ox} is the SiO_2 thickness.

Fig. 2 shows a V_{FB} versus SiO_2 thickness plot for TiN/SiO_2/p-Si MOS structure. The experimental data points of Fig. 2 are fit to a straight line using "least square fit". The intercept of the straight line is the value of φ_{ms} for TiN metal gate electrode. To obtain the TiN work function (φ_m), the following equation was used

$$\varphi_m = \varphi_{ms} + \varphi_s = \varphi_{ms} + \left(\chi + \frac{E_g}{2q}\right) + \left(\frac{kT}{q} \ln\left(\frac{N_a}{n_i}\right)\right), \tag{2}$$

where χ is the Si electron affinity (4.05 eV), E_g is the band gap of the Si (1.12 eV), n_i is the intrinsic carrier concentration in Si and Na is the channel doping levels (~ 10^{18} cm^{-3}). In this study, the φ_m for TiN metal gate electrode is found to be 4.31, which is consistent with the reference [10].

However, the values of φ_m for TiN increase with increasing of the N_2 partial pressure during deposition process. To achieve higher φ_m value, the diffusion of N_2 towards the dielectric/electrode interface is necessary as it is produced in a more stoichiometric TiN metal gate electrode [10].

When characterizing high-k gate stacks, it is important to be aware of the charges in the different interfaces and layers that the stack comprises. The dielectric layers in the high-k gate stacks consist of the bi-layer structure, as shown in Fig. 1. Different types of charges are located at the different dielectric layer as well as interface.

The fundamental equation that relates the V_{FB} to the gate dielectric charge distribution per volume, $\rho(x)$, φ_{ms} and EOT of the MOS stack structure, can be expressed as,

$$V_{FB} = \varphi_{ms} - \frac{1}{\varepsilon_{SiO_2}} \left[\int_0^{EOT} x\, \rho(x) dx \right] \tag{3}$$

Considering the bi-layer stack structure [11], the equation (3) can be rewritten as,

$$V_{FB} = \varphi_{MS} - \frac{1}{\varepsilon_{sio_2}} \left[\int_0^{EOT_{hk}} x\, \rho(x) dx \right] - \frac{1}{\varepsilon_{OX}} \left[Q_{int\, IL}\, EOT \right] - \\ - \frac{1}{\varepsilon_{OX}} \left[\frac{1}{2} (\rho_{bulkhk} EOT)^2 \right] + \frac{1}{\varepsilon_{OX}} \left[\frac{1}{2} (\rho_{bulkhk} EOT)^2 \right] \tag{4}$$

to Contents, page 1

If EOT_{hk} is fixed and the total EOT (EOT = EOT_{hk} + EOT_{IL}) is varied only by changing EOT_{IL},one can get the expression for V_{FB} with EOT as a polynomial of order two [11].

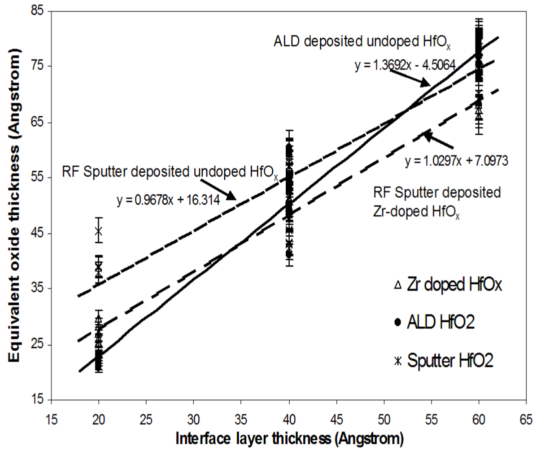


Fig. 3 *– Plot of EOT versus the interface layer physical thickness to determine the physical thickness of high-k layer for different stack layer*

Based on recent reports [11-13], the bulk charge in both layers is usually much less than the interface charge (i.e. $\rho_{bulk\ \mathrm{hk}}.\mathrm{EOT}_{hk} << Q_{int\ hk}$). Therefore, the equation 4 can be simplified and V_{FB} can be expressed in terms of EOT,

$$V_{FB} = \varphi_{ms} - \frac{1}{\varepsilon_{\mathrm{SiO_2}}}\left[Q_{\mathrm{int}\,hk} \cdot \mathrm{EOT}_{hk}\right] - \frac{1}{\varepsilon_{\mathrm{SiO_2}}}\left[Q_{\mathrm{int}\,IL}\right] \cdot \mathrm{EOT} \tag{5}$$

To find the values of $Q_{int\ IL}$ and $Q_{int\ hk}$ charges present in the different location of the stacked dielectrics (as shown in Fig. 1), the φ_{ms} and EOT_{hk} can be extracted properly. According to equation (1), the Y-axis interception in the V_{FB}-SiO_2 thickness plot represents φ_{ms}, while according to equation (5); the Y-axis interception in the V_{FB}-EOT plot represents the φ_{ms} plus the effect of the HfO_2/SiO_2 interface charges. We have calculated the φ_{ms}values using the TiN/SiO_2/p-Si MOS capacitor system and the equation (1), as there is no interfacial layer. However to calculate EOT_{hk}, we can use the following equation,

$$\mathrm{EOT} = \frac{\varepsilon_{\mathrm{SiO_2}}}{\varepsilon_{IL}} d_{IL} + \frac{\varepsilon_{\mathrm{SiO_2}}}{\varepsilon_{hk}} d_{hk} \tag{6}$$

where the first term is the contribution of EOT for SiO_2 rich dielectric layer and second term is the EOT for high-k dielectric layer. The ε_{IL}, ε_{hk}, d_{IL} and d_{hk} are the dielectric constant of interfacial layer, high-k layer, the physical thickness of the interfacial layer and the physical thickness of the high-k layer, respectively. We have considered the ε_{IL} is the dielectric constant of SiO_2 rich layer and we have used the fixed Zr-doped HfO_x and undoped HfO_x

to Contents, page 1

film on terraced oxide samples. The EOT_{hk} value for doped and undoped films can be calculated from the intercept on the EOT axis when EOT is plotted against the SiO_2 thicknesses, as shown Fig. 3.

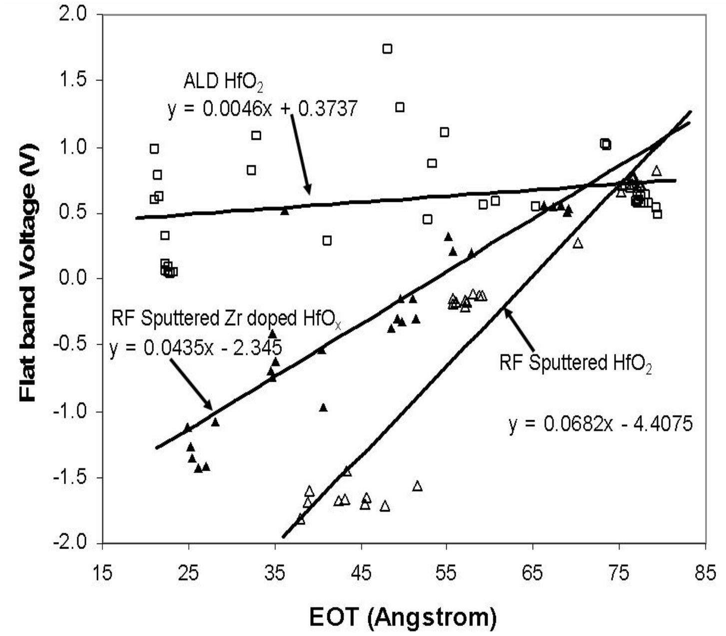


Fig. 4 *– Plot of flat band voltage versus EOT (by changing the Interfacial layer thickness) for determination of the different charges located at different interfaces using ALD and RF-sputter deposited HfO_2, and RF-sputter deposited Zr doped HfO_x with TiN gate electrode MOS capacitor structure*

After knowing the values of φ_{ms} (from equation 1) and EOT_{hk} (from equation 6), one can extract the values of $Q_{int\ hk}$ and $Q_{int\ IL}$ using equation 5 for stacked MOS capacitor with undoped HfO_2 and Zr-doped HfO_2 (shown in Table 1) films.

From Fig. 2, it is observed that the V_{FB} has shifted toward the positive direction as the SiO_2 thickness increases supports that the negative Q is primarily located in the SiO_2 near to the Si/SiO_2 interface. The positive drift of V_{FB} with increasing SiO_2 thickness was caused by negative electric centers. The probable causes to form the negative electric centers are the trapping of electrons by the unsaturated bonds, e.g. Si-O in the SiO_x film and absorbed impurities (i.e. Na^+, K^+) [14]. The $SiO_4{}^{4-}$ tetrahedral network is the basic structure of SiO_2 film despite whether it is crystalline or amorphous. There are always unsaturated bonds of oxygen in the surface layer and interfacial layer of SiO_2/SiO_2. Thermal treatment causes continuous oxygen diffusion from the surface to the SiO_2/SiO_2 interface. Hence, rich oxygen anions accumulate in these surface and interfacial layer, forming negative electric centers. From Table 1, it is seen that as long as the high-k film has an inserted thermal SiO_2 interface, the interfacial charges near SiO_x/Si interface are negative, same as $TiN/SiO_2/Si$ system independent of the high-k film's deposition method or its doping level.

to Contents, page 1

Table 1 *– Comparison of different charges located in the high-k/SiO_2 interface and Si/SiO_2 interface for different gate dielectrics are tabulated here*

Sample	$Q_{int\ hk}$ (cm^{-2})	$Q_{int\ IL}$ (cm^{-2})
TiN//SiO_2/p-Si	---	$-5.24 \cdot 10^{12}$
TiN/ALD deposited HfO_2/SiO_2/p-Si	$-5.71 \cdot 10^{13}$	$-9.92 \cdot 10^{11}$
TiN/RF-Sputter deposited Zr-doped HfO_2/SiO_2/p-Si	$4.6 \cdot 10^{13}$	$-9.38 \cdot 10^{12}$
TiN/RF-Sputter deposited HfO_2/SiO_2/p-Si	$4.7 \cdot 10^{13}$	$-1.29 \cdot 10^{13}$

with EOT for ALD deposited undoped HfO_2 on SiO_2/Si samples compared to the other samples, e.g. thermally grown SiO_2 on Si samples, and sputter deposited Zr-doped and undoped HfO_2 gate dielectrics on SiO_2/Si samples. This can be explained as the formation of more negative interfacial charges in the high-k/SiO_2 interface for ALD deposited samples. However, the higher slope value is observed for sputter deposited undoped HfO_2 samples (in Fig. 4), which signifies that more positive $Q_{int\ hk}$ are present compared to Zr doped HfO_x samples also shown in Table1. The existence of positive $Q_{int\ hk}$ can be explained considering the gate sputtering process. The sputtering process can cause two kinds of damages: the surface damage and bulk damage [15]. The surface damage is mainly caused by the ion bombardment. The plasma radiation is responsible for bulk damage of the dielectrics. The plasma radiation, for example, UV and the high energy photons could create traps in the gate dielectrics or at the interface. These damages may generate lots of positive interface charges near the high-k/SiO_x interface. Recently, Tewg et al. has also reported the existence of positive charge defects in Zr doped TaO_x dielectric films [16]. The high negative charge values of $Q_{int\ hk}$ for ALD deposited undoped HfO_2 samples agree with the literature reports [13, 17]. Again from Table 1, it is seen that the interface charges present in both high-k/SiO_x and SiO_x/Si interface are less due to insertion of the Zr atoms in HfO_x gate dielectrics which is due to the reduction of dangling or unsaturated bonds of excess oxygen in HfO_x [5].

4. CONCLUSIONS

We have found out the metal gate work function value (4.31 eV) for TiN gate electrode using the TiN/SiO_2/p-Si capacitor structure. The different types of charges present in the gate dielectrics are reported here. The polarity of interfacial charges at the Si/SiO_2 interface is the same for SiO_2, ALD and RF sputter deposited undoped HfO_2 and also RF sputter Zr doped HfO_x samples. The increasing slope of the V_{FB} versus EOT plot (in Fig. 2 and 4) indicate the shifting from negative to positive charge quantities present at the high-k/SiO_2 interface. The influence of Zr doping in HfO_x gate dielectrics on the interfacial charge defects found the suitability of the doping element in HfO_x dielectrics for future CMOS generation.

This research was partially supported by NSF project DMI-0300032 and SEMATECH project 309366.

to Contents, page 1

REFERENCES

1. C.K. Maiti, S.K. Samanta, S. Chatterjee, G.K. Dalapati, L.K. Bera, *Solid-State Electron.* **48,** 1369, (2004).
2. S. Chatterjee, S.K. Samanta, H.D. Banerjee, C.K. Maiti, *Thin Solid Films* **422,** 33 (2002).
3. J.-Y. Tewg, Y. Kuo, J. Lu *Electrochem. Solid St.* **8,** G27, (2005).
4. Y. Kuo, J. Lu, J. Yan, S. Chatterjee, T. Yuan, H.C. Kim, J. Peterson, M. Gardner, *208th meeting of the Electrochemical. Society, Abs. # 548*, (Los Angeles, California, USA: 2005).
5. L. Manchanda, M.D. Morris, M.L. Green, R.B. van Dover, F. Klemens, T.W. Sorsch, P.J. Silverman, G. Wilk, B. Busch, S Aravamudhan, *Microelectron. Eng*. **59,** 351 (2001).
6. I. De, D. Johri, A. Srivastava, C.M. Osburn, *Solid-State Electron*. **44,** 1077 (2000).
7. Y. Yeo, P. Ranade, T. King, C. Hu, *IEEE Electron Dev. Lett*. **23,** 342 (2002).
8. G.A. Brown, G. Smith, J. Saulters, K. Matthews, H.C. Wen, P. Majhi, B.H. Lee, IEEE SISC (San Diego: 2004).
9. J.R. Hauser, K. Ahmed, *AIP Conf. Proc*. **449,** 235 (1998).
10. J. Westlinder, G. Sjoblom, J. Olsson, *Microelectron. Eng*. **75,** 389 (2004).
11. R. Jha, J. Gurganos, Y.H. Kim, R. Choi, J.C. Lee, V. Misra, *IEEE Electron Dev. Lett*. **25,** 420 (2004).
12. A. Kerber, E. Cartier, L. Pantisano, R. Degraeve, T. Kauerauf, Y. Kim, A. Hou, G. Groeseneken, H.E. Maes, U. Schwalke, *IEEE Electron Dev. Lett*. **24,** 87 (2003).
13. G.D. Wilk, M.L. Green, M.-Y. Ho, B.W. Busch, T.W. Sorsch, F.P. Klemens, B. Brijs, R.B. van Dover, A. Kornblit, T. Gustafsson, E. Garfunkel, S. Hillenius, D. Monroe, P. Kalavade, J.M. Hergenrother, *IEEE VLSI Symp. Tech. Dig*. 88 (2002).
14. L.K. Bera, W.K. Choi, C. S. Tan, S.K. Samanta, C.K. Maiti, *IEEE Electron Dev. Lett*. **22,** 387 (2001).
15. Y. Kuo, *Appl. Phys. Lett*. **61,** 2790 (1992).
16. J.-Y. Tewg, Y. Kuo, J. Lu, B.W. Schueler, *J. Electrochem. Soc*. **152,** G617 (2005).
17. Z. Zhang, M. Li, S.A. Campbell, *IEEE Electron Dev. Lett*. **26,** 20 (2005).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.170-178

PACS numbers: 61.46.Hk, 81.07.Bc

OXIDE NANOSTRUCTURES: CHARACTERIZATIONS AND OPTICAL BANDGAP EVALUATIONS OF COBALT-MANGANESE AND NICKEL-MANGANESE AT DIFFERENT TEMPERATURES

***C.R. Indulal*[1]*, G.S. Kumar*[1]*, A.V. Vaidyan*[2]*, R. Raveendran*[1]**

[1] Department of Physics,
S N College, Kollam, India
E-mail: cr_indulal@yahoo.com

[2] Department of Chemistry,
St Stephens College, Pathanapuram, India

Cobalt-Manganese and Nickel-Manganese oxide (CoMnO and NiMnO) nanoparticles were prepared by chemical co-precipitation method by decomposition of their respective metal sulfides and sodium carbonate using ethylene diamene tetra acetic acid as the capping agent. The samples were heated at 400, 600 *and* 800 *°C. The average particle sizes were determined from the X-ray line broadening. The diffractograms were compared with JCPDS data to identify the crystallographic phase and cubic structure of the particles. The samples were characterized by XRD, FTIR and UV analyses. The internal elastic micro strains were calculated and it was seen that as the particle size increases strain decreases. The FTIR studies have been used to confirm the metal oxide formation. The chemical compositions of the samples were verified using EDX spectra. The surface morphologies of the samples were studied from the SEM images. The absorption spectra of the materials in the UV-Vis-NIR range were recorded. From the analysis of the absorption spectra, the direct band gaps of the materials were calculated.*

Keywords: *NANOPARTICLES, ARRESTED PRECIPITATION, SEM, EDX, OPTICAL BAND GAP.*

(Received 04 February 2011, in final form 18 March 2011)

1. INTRODUCTION

Recently nano oxide particles have been drawing much attention because of their peculiar optical [1, 2] and magnetic [3] properties. Due to the large surface to volume ratio of nano particles, their properties (electrical, optical, chemical, mechanical and magnetic) can be selectively controlled by engineering the size, morphology and composition of the particles. Nano crystalline materials exhibiting this large surface area can be applied to gas sensors for which an excellent surface effect is required. These new assemblies can have enhanced or entirely different properties from their parent bulk materials [4].The fine particles of the compounds with very small size exhibits unique UV absorbing ability, high stability at high temperatures, high hardness and reactivity as catalyst [5, 6].

Nanoparticles of Cobalt-Manganese oxide and Nickel-Manganese oxide were prepared by chemical co-precipitation method. Doped Cobalt and Nickel oxides show P-type semi conducting behaviour similar to intrinsic spinel Cobalt oxide. Cobalt-Manganese oxide and Nickel-Manganese oxide systems have gained immense importance because of their potential applications such as electrodes in batteries; in solar cells, in super capacitor, in sensors and in switches.

to Contents, page 1

2. EXPERIMENTAL PROCEDURE

Cobalt-Manganese and Nickel-Manganese oxide nanoparticles were prepared by chemical co-precipitation method from analytical grade cobalt sulphate (0.4M), nickel sulphate (0.4M), manganese sulphate (0.4M) and sodium carbonate (0.6M) using ethylene diamene tetra acetic acid as the capping agent. The metal carbonate precipitates were separated from the reaction mixture and washed several times with distilled water and then with ethanol to remove the impurities and traces of EDTA and original reactants if any. The wet precipitates were dried and thoroughly ground using an agate mortar to obtain the metal carbonate precursor in the form of fine powder. On heating to the required temperatures (400, 600 and 800 °C) the metal carbonate precursor decomposes to metal oxides.

2.1 Characterization of the sample

X-ray diffraction is an ideal technique for the determination of crystallite size of the powder samples. The basic principle for such a determination involves precise quantification of the broadening of the diffraction peaks. Based on this principle, a few techniques involving Scherrer equation, integral breadth analysis or Hall-Williamson approach and Fourier method of Warren-Averbach have been developed [7-9]. XRD studies were carried out using XPERT-PRO powder diffractometer (PAN analytical, Netherlands) employing Cu- K_{α} radiation in the 2θ range 10° to 70° at 30 mA, 40 kV. The TGA/DTA of the nano particles were taken using Perkin-Elmer, Diamond TG/DTA apparatus. The morphologies of the powder samples were characterized by scanning electron microscope (SEM) JEOL/EO JSM-6390. The energy dispersive analyses of X-rays (EDX) were carried out on the samples to ascertain the composition. The infrared spectroscopic (IR) studies of the oxides were made using Perkin-Elmer FTIR spectrophoto meter in the wavenumber range 500 and 4000cm^{-1} by KBr disc method. The UV spectrum were obtained using Shimadzu UV-2550 UV visible spectrophoto meter.

Table 1 *– Table 1a and 1b, variation of particle sizes, elastic strains and optical band gaps of CoMnO and NiMnO nanoparticles prepared at different sintering temperatures*

	Sintering temperature	Particle size (nm)	Elastic strains	Bandgap (eV)
a	400°C	10	0.002305	1.42
	600°C	26	0.000918	1.38
	800°C	31	0.000752	1.30

	Sintering temperature	Particle size (nm)	Elastic strains	Bandgap (eV)
b	400°C	8	0.002620	1.44
	600°C	17	0.001313	1.41
	800°C	41	0.000562	1.37

to Contents, page 1

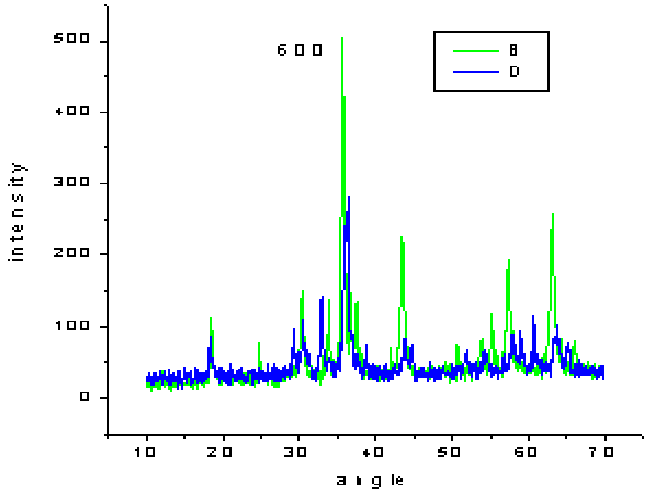


Fig. 1 *– XRD pattern of the CoMnO and NiMnO sintered at* 600 *°C*

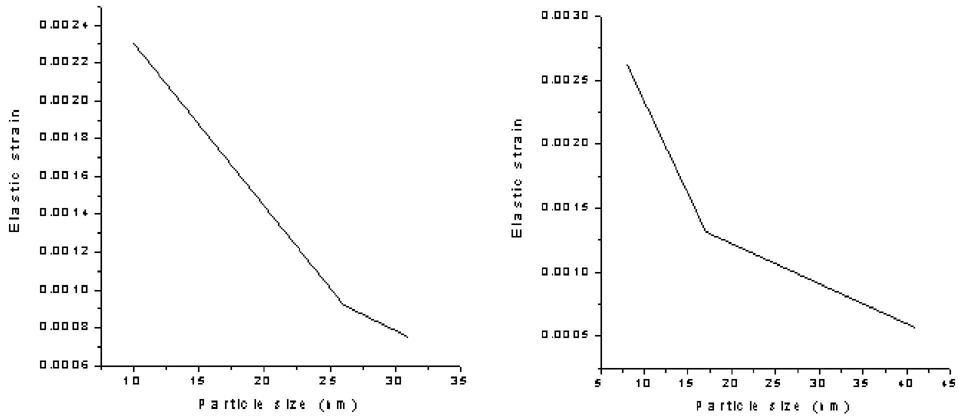


Fig. 2 *– Graph showing variation of elastic strain with particle size for CoMnO (a) and NiMnO (b) nanoparticles*

3. RESULTS AND DISCUSSIONS

3.1 XRD Studies

XRD pattern reveal that the particles are nano sized and crystalline. The particle sizes are calculated from Scherrer equation, $d = 0.9\lambda/\beta\cos\theta$ [10], where β represents the full width at half maximum (FWHM) of XRD lines, $\lambda = 1.54060$ Å. The XRD patterns of CoMnO and NiMnO sintered at 600 °C are shown in Fig. 1. The most intense peaks (intensity 100) are from the (311) planes. The crystallite sizes of CoMnO and NiMnO at various calcination temperatures (400, 600 and 800 °C) are as shown in Table 1a and 1b respectively. As the sintering temperature increases, the particle size increases. This indicates that the size of the crystallites can be adjusted by

to Contents, page 1

controlling the temperature of the reaction [11]. The diffractogram was compared with JCPDS (File No 23-1237 and 01-1110) data to identify the crystallographic phase and cubic structure of the particles. The XRD pattern when compared with JCPDS reveals the structures as spinel oxides. The broadening of the peaks in the XRD pattern may be due to the micro straining of the crystal structures arising from the defects like dislocations and twinning. These are believed to be associated with chemically synthesized nanocrystals. As the crystals grow spontaneously during chemical reaction, the legands get negligible time to diffuse to an energetically favorable site resulting in large crystal defects [12].

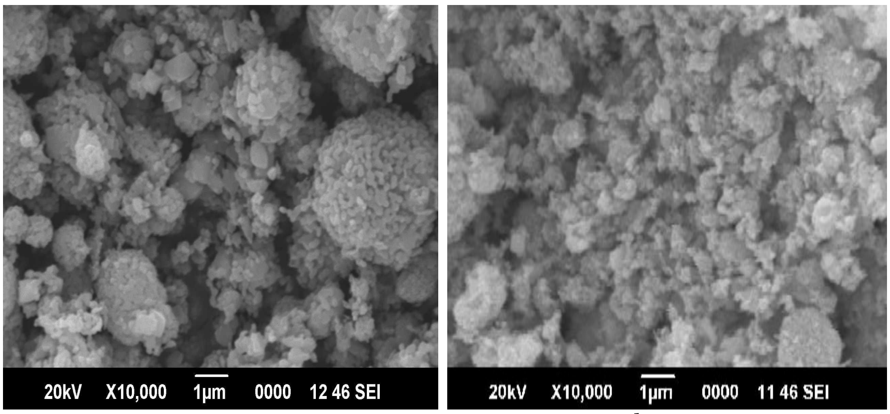
The elastic strain of the materials can be calculated using the formula $E = \beta/2 \cot \theta$ [13]. The variation of the elastic strain with particle size is shown in the Fig.2a and 2b. It can be seen that as the particle size increases elastic strain decreases. This strain contributes to the broadening of the XRD pattern.

3.2 Thermal analysis

Thermo gravimetric analysis of the carbonate precursor was carried out to determine the decomposition temperature and the rate of decomposition. The decomposition temperature is found to lie between 300 and 350 °C. Thus the heat treatment of the ground precursor powders at their respective decomposition temperature and beyond results in the evolution of heat from the combustion of the residual carbonaceous material. This facilitates the reaction among the constituent metal ions and the formation of the desired oxide phase at relatively low external temperature.

3.3 Micro structural studies

For the micro structural studies, the samples heated at 600°C were directly transferred in to the chamber of the SEM with out disturbing the original nature of the products. SEM images and the energy dispersive spectra (EDX) of CoMnO and NiMnO are shown in the Fig. 3a, 3b and 4a, 4b respectively. The SEM pictures reveal that the particles are more or less spherical in shape.



a b

Fig. 3 – SEM image of CoMnO (a) and NiMnO (b)

to Contents, page 1

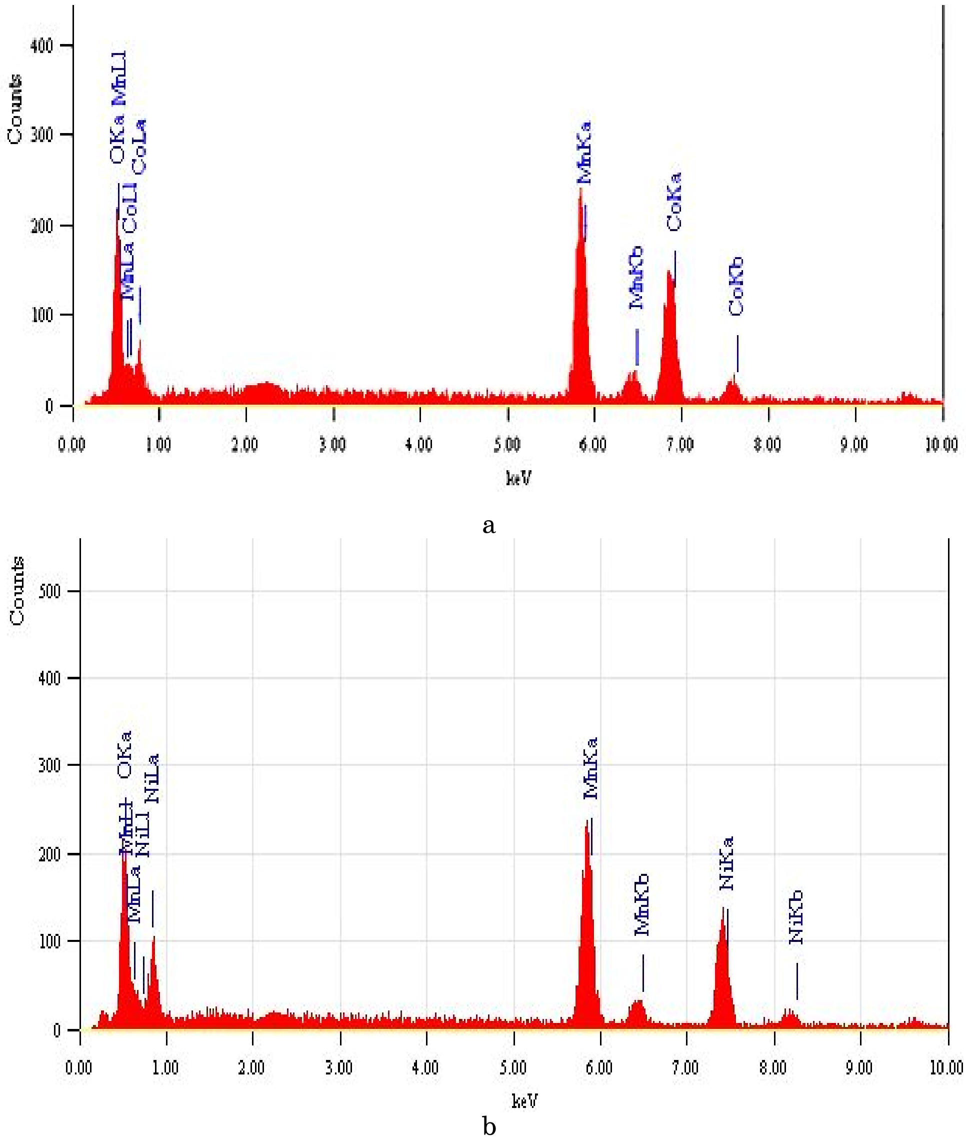
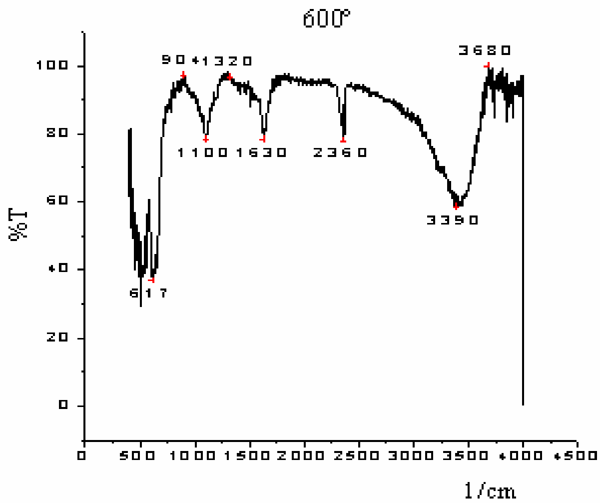


Fig. 4 *– EDX spectrum of CoMnO (a) and NiMnO (b)*

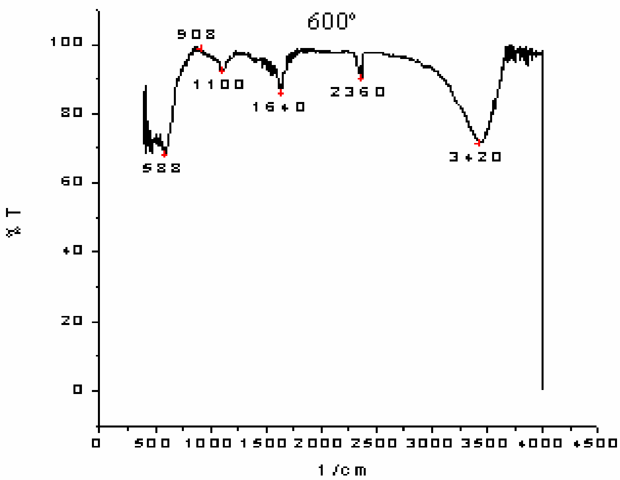
3.4 FTIR Spectra analysis

The FTIR spectra of the samples are shown in Fig. 5a and 5b respectively. The broad absorption bands in the region around 3390 and 3420 cm^{-1} are due to the presence of co-ordinated / entrapped water. The presence of some carbonaceous materials is evident from the IR spectra which depicts the bands at 1630 and 1640 cm^{-1} corresponding to carboxylate ions [14, 15]. The band around 600cm^{-1} may be due to the bending modes of the metal oxides. The other bands may be due to the micro structural formation of the samples.

to Contents, page 1



a



b

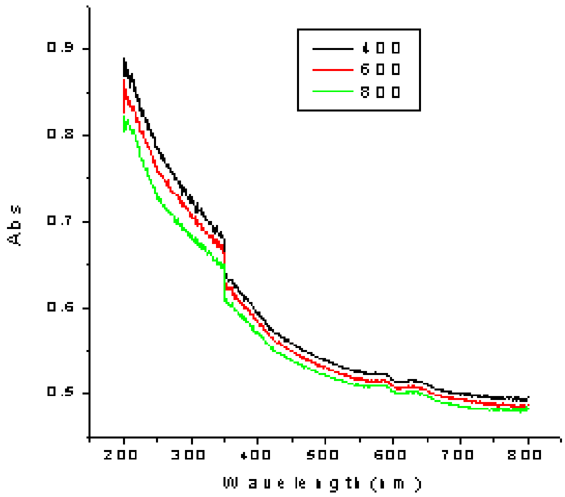
***Fig. 5** – FTIR spectrum of CoMnO (a) and NiMnO (b)*

3.5 UV spectral studies

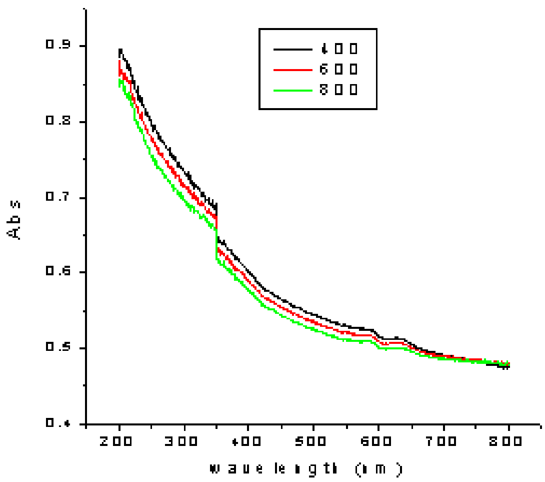
The UV spectra of CoMnO and NiMnO sintered at 400, 600 and 800 °C are taken in the wavelength range of 200 to 800 nm with 1 nm resolution is shown in fig 6a and 6b respectively. UV spectra provide important information about the details related with optical band gap of the material. The energy band of the material is related to the absorption coefficient α by the Tauc relation, $\alpha h\nu = A(h\nu - E_g)^n$, where A is a constant, hυ is the photon energy ($\nu = c/\lambda$), Eg is the band gap and $n = 1/2$ for an allowed direct transition. Plotting a graph between $(\alpha h\nu)^2$ and $h\nu$ (Fig. 7) gives the value of direct band gap of CoMnO and NiMnO sintered at 600 °C [16]. The

to Contents, page 1

extrapolation of the straight line to $(\alpha h\nu)^2 = 0$, gives the value of the band gap. From the UV spectra, it is clear that the absorbance decreases with increase in wavelength. This decrease in the absorption indicates the presence of optical band gap in the material. This corresponds to the excitation of surface plasmons in the composite nano particles. The optical band gap of the materials determined from the absorption spectra is shown in the Table 1. As the sintering temperature increases, the particle size increases and the band gap decreases.

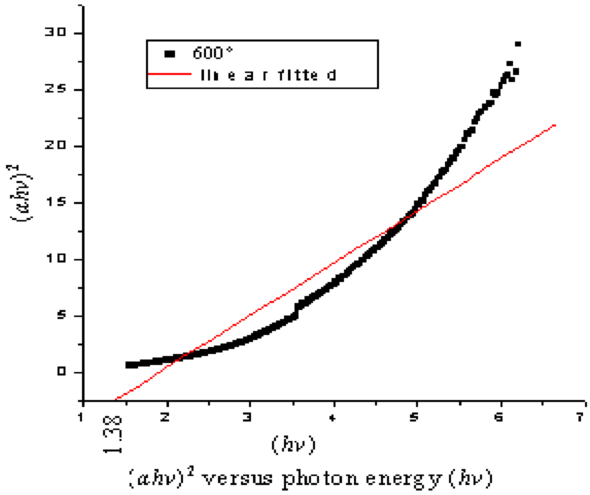


a

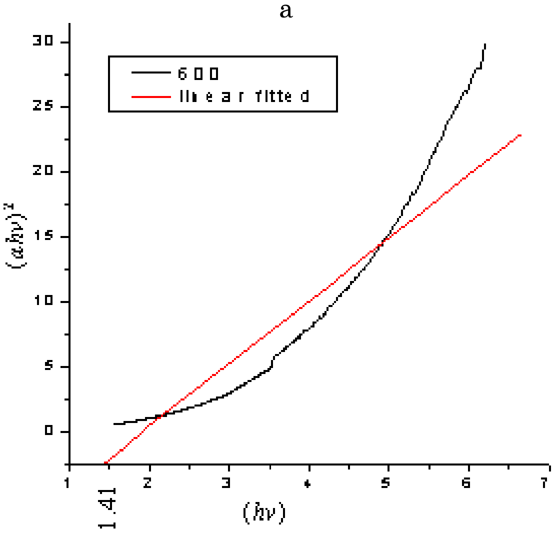


b

Fig. 6 – *Absorbance versus wavelength graphs for CoMnO (a) and NiMnO (b) at* 400, 600 *and* 800 *°C*



a



b

Fig. 7 *–* $(\alpha h\nu)^2$ *versus hυ graphs for the samples sintered at* 600 *°C*

REFERENCES

1. R. Lopez, T.E. Haynes, L.A. Boatner and R.F. Haglund, *Opt. Lett.* **27,** 1327 (2002).
2. M. Ikeyama, S. Nakao, M. Tazawa, K. Kamada, *Nucl. Instrum. Meth. B* **175,** 652 (2001).
3. M. Muroi, R. Street, P.G. McCormick, *J. Appl. Phys.* **87,** 3424 (2000).
4. Kenneth J Klabunde & Cathymohs, Department of Chemistry, Kansas state university, *Manhanttan Kansas second Edition*: 1998).
5. M.S. Tsai, *Mater. Sci. Eng. B.* **110,** 132 (2004).
6. T. Mori, Y.R. Wang, J. Drennan, *Solid State Ionics* **175,** 641 (2004).
7. P. Scherrer, *Math. Phys. K.* **1**, 98 (1918).

to Contents, page 1

8. G.K. Williamson, W.H. Hall, *Acta Metall.* **1**, 22 (1953).
9. B.E. Warren, B.L. Averbach, *J. Appl. Phys.* **21**, 595 (1950).
10. B.D. Cullity, *Elements of X-ray diffraction*, (Addison-Wesley publishing company, Inc. California: 1970).
11. L.P. Wang, G.Y. Hong, S.C. Qu, Z.G. Wang, *Nanotechnology* **16**, 1469 (2005).
12. H.C. Warad, S.C. Ghosh, Hemtanon, *Sci. Technol. Adv. Mater.* **6**, 296 (2005).
13. R. Kelsall, I. Hamley, M. Geoghegan, *Nano Scale Science & Technology*, (John Wiley: Chichester: 2005).
14. G. Ehrlich, *J. Am. Chem. Soc.* **76**, 5263 (1954).
15. F. Douville, G. Duval & J. Lecomte, *Compt. Rend. Acad. Sci.* **212**, 963 (1941).
16. D. Mohanta, S.S. Nath, N.C. Mishra, A. Choudhury, *Bull. Mater. Sci.* **26**, 289 (2003).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.179-184

PACS numbers: 71.20.Be, 68.37.Og

SYNTHESIS AND CHARACTERIZATION OF NOVEL NANOCRYSTALLINE ZIRCONIUM (IV) TUNGSTATE SEMICONDUCTOR

S. Manoj, B. Beena

Department of Chemistry, D.B. College,
Sasthamcotta, Kollam, Kerala, India 690 521
E-mail: beenadbc@gmail.com

Nanocrystalline zirconium (IV) tungstate is prepared by chemical coprecipitation method using ethylene diamine tetra acetic acid as the templating agent. Elemental composition is determined by EDS. The characteristic bonding position is identified using FTIR.XRD is used to find the theoritical value of size and phase identification using JCPDS. Morphology is examined using SEM and HRTEM. UV absorption at 260 *nm corresponds to an energy gap of* 4.48 *eV, characteristic of semiconducting nanoparticles.*

Keywords: *ZIRCONIUM (IV) TUNGSTATE, SEMICONDUCTOR, BAND GAP, EDS, HRTEM, ZIRCONIA, TUNGSTEN OXIDE*

(Received 04 February 2011, in final form 20 March 2011)

1. INTRODUCTION

Metal oxide semiconductor nanomaterials are usable as semiconducting gas sensors because they have many advantage such as small diamensions, low cost, low power consumption, on-line operation and high compatibility with microelectronic processing. These include SnO_2, WO_3, ZrO_2, WO_3, etc. These materials have non-stoichiometric structures, so free electrons orginating from oxygen vacancies contribute to electronic conductivity [1]. Tungsten oxide is an n-type semiconductor with energy gap in the range of 5.59 - 5.7 eV. Therefore, it is suitable for applications such as electrochromic, optochromic and gasochromic coatings for smart windows, information display and various sensors [2]. Many researchers have studied the sensor, magnetic and electric properties of WO_3 prepared by heat treatment, arc discharge, sol gel etc with different morphology and structure. These materals are studied in different forms such as nanowires, nanorods, dopped with transition metal, etc [3, 4, 5, 6] which are either in the isolated form or in mixed oxide form. These mixed oxide forms are non-stoichiometric forms with varying metal concentration.

In the present work, nanoform of zirconium (IV) tungstate is prepared in the stoichiometric form by chemical coprecipitation method. The material is well characterized using EDS, XRD, FTIR, SEM and HRTEM. UV spectrum is used to find the optical band gap.

to Contents, page 1

2. EXPERIMENTAL

2.1 Preparation of zirconium (IV) tungstate in the nanoform

The nanoform of zirconium (IV) tungstate is synthesized by chemical coprecipitation method. In this, equimolar aqueous solutions of zirconium oxychloride and sodium tungstate in the volume ratio 1:2 is added to ethylene diamine tetra acetic acid template at a pH ~ 1.5 with constant stirring using a magnetic stirrer. The precipitate of zirconium (IV) tungstate is separated by filtration and washed with distilled water. It is then converted to the hydrogen form by immersing in 1M HCl.

2.2 Characterization of the material

EDS spectrum of the material was taken using JOEL Model JED-2300. X-ray diffractogram ($2\theta = 10$ - $90°$) was obtained on XPERT-PRO powder diffractometer with Cu-Kα radiation. The FTIR spectrum was recorded using KBr wafer on the Thermo Nicolet FTIR model AVATAR 370 DTGS.SEM was taken from BJOEL Model JSM-6390LV.PHILIPS Model CM 200 was used to take the HRTEM. The absorption spectra were recorded at room temperature using SHIMADSU UV-2550 UV Visible spectrophoto-meter.

2.3 Results and Discussion

EDS spectrum shows the presence of zirconium and tungsten in the stichiometric proportion. The empirical formula calculated is $Zr_2W_3O_{13}$. The EDS spectrum is shown in figure 1. The X-ray diffractions of the nanoparticles of zirconium (IV) tungstate consist (figure 2) of cubic phases of ZrW_2O_8, agreeing with the JCPDS file No.89-6670 with small amount of isolated cubic zirconia (JCPDS file No.89-9069) and hexagonal tungsten oxide (JCPDS file No.82-2459).

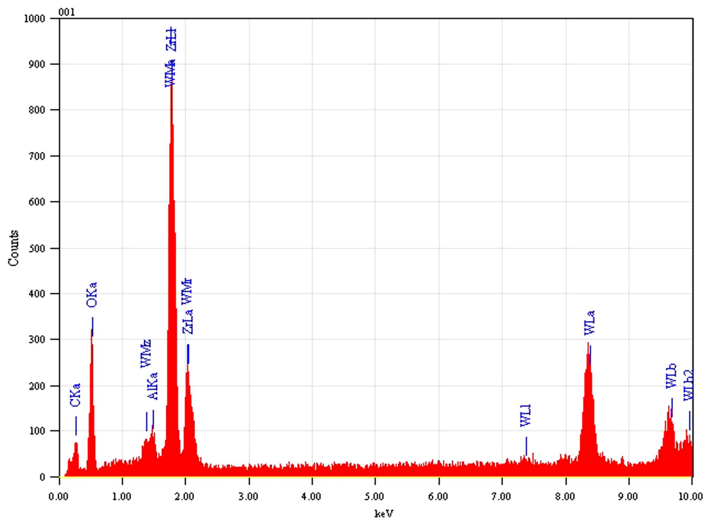


Fig. 1 – EDS of zirconium (IV) tungstate

to Contents, page 1

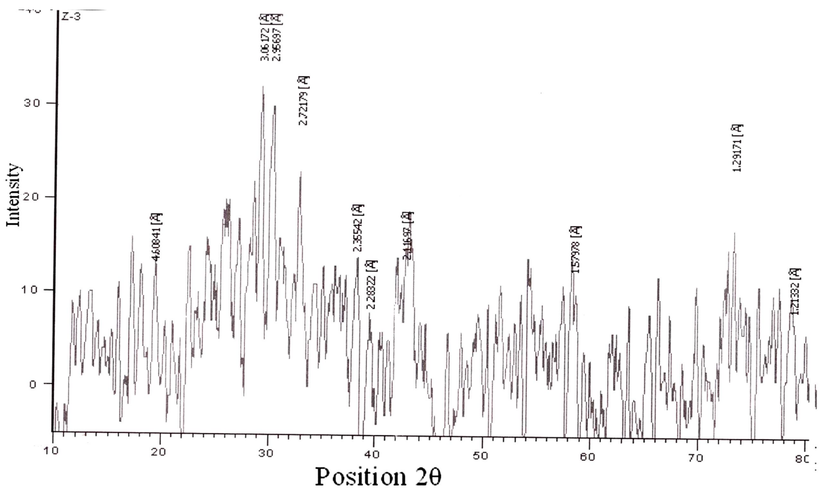


***Fig. 2** – XRD of zirconium (IV) tungstate*

FTIR spectrum (figure 3) shows bands due to water molecules/surface hydroxyl groups (~ 3415 cm $^{-1}$, 1637 cm $^{-1}$, 1380 cm $^{-1}$). Band in the region 837 cm is due to WO_4^2 species and at 711 cm $^{-1}$ is due to W-O stretching vibration. Band in the region 493 cm $^{-1}$ is due to Zr-O vibration [7, 8].

SEM and HRTEM shown in figure 4 and figure 5 are used to examine the morphology of the nanomaterial. The spherical nature of the particles is revealed from both the SEM and HRTEM. Further from HRTEM it is seen that the particles are less than 50 nm in size.

UV Spectrum shown in figure 6 gives information about excitonic or interband transition of nanocrystalline material. It shows absorption at 260 nm which corresponds to an energy gap of 4.77 eV. The fundamental absorption which corresponds to electron excitation from the valence band to the conduction band is used to determine the nature and the value of optical band gap. The relation between absoption coefficient (α) and incident photon energy ($h\nu$) is given by the Tauc relation,

$$(\alpha h\nu)^{1/n} = A(h\nu - E_g),$$

where A is a constant, E_g is the band gap of the material, α, is the absorption coefficient and the exponent ‘n’ depends on the type of transition [9]. The value of band gap can be determined by plotting $(\alpha h\nu)^{1/n}$ versus $h\nu$ (figure 7) and extrapolating the straight line portion of the graph to $h\nu$ axis. The direct allowed band gap calculated is 4.48 eV. A more straight forward method of determining an approximate value of band gap regardless of band structure is by plotting the absorbance versus energy, and extrapolating to zero absorbance which in this case gives a band gap of 4.49 eV. Bulk ZrW_2O_8 is an insulating dielectric ceramic and the calculated band gap for α-phase is 2.81 eV and for γ-phase is 1.74 eV [10]. The increase in the band gap to 4.48 eV of this material is due to small size and the quantum confinement limiting. The quantum confinement effect is expected for semiconducting nanoparticles and the absorption edge will be shifted to high energy with decrease in particle size [11].

to Contents, page 1

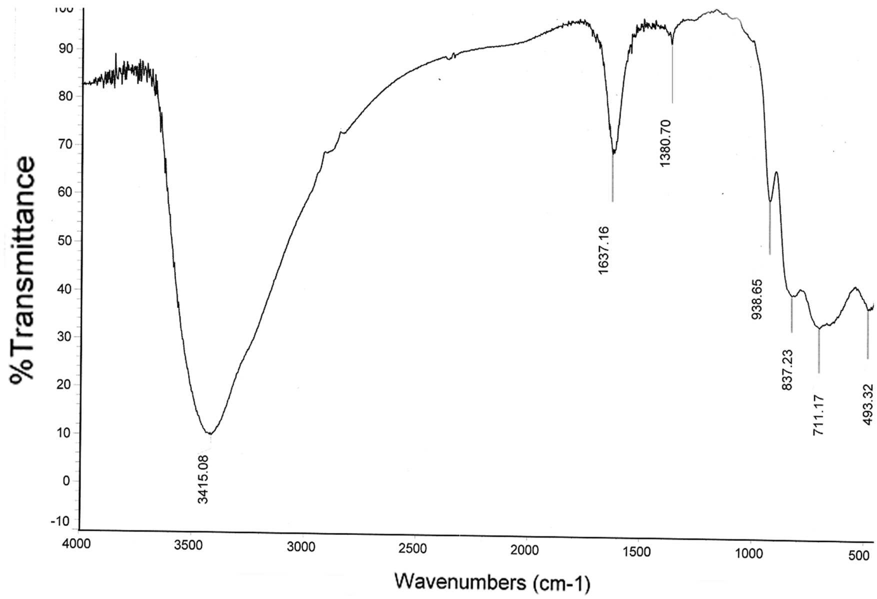


Fig. 3 – *FTIR of zirconium (IV) tungstate*

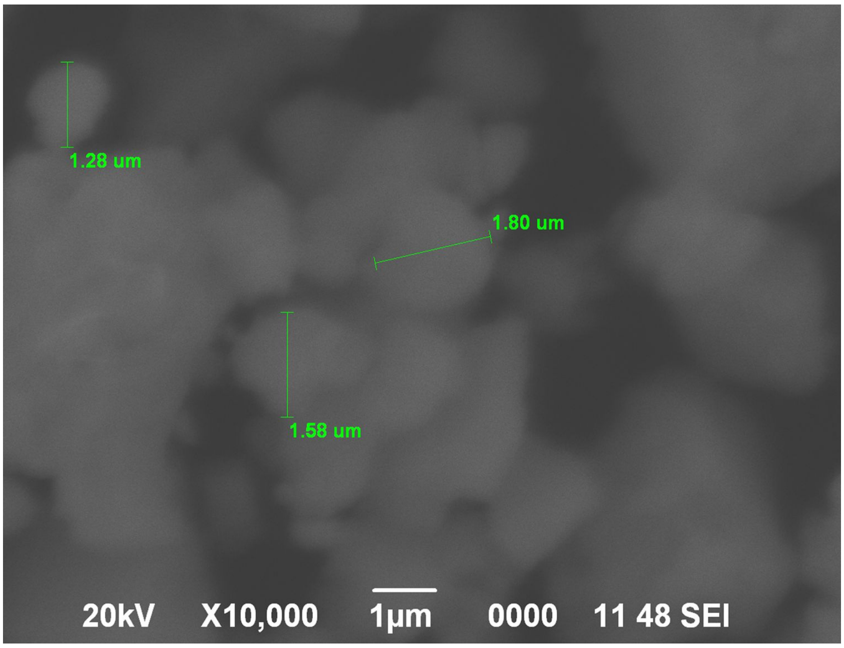


Fig. 4 – *SEM of zirconium (IV) tungstate*

to Contents, page 1

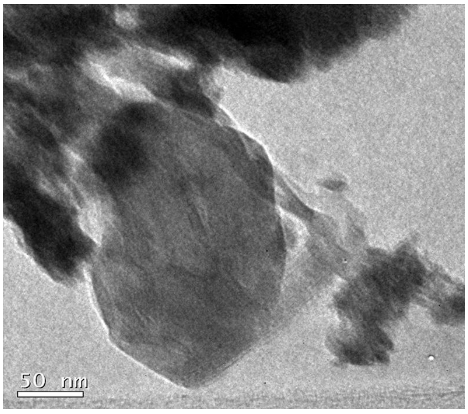


Fig. 5 – *HRTEM of zirconium (IV) tungstate*

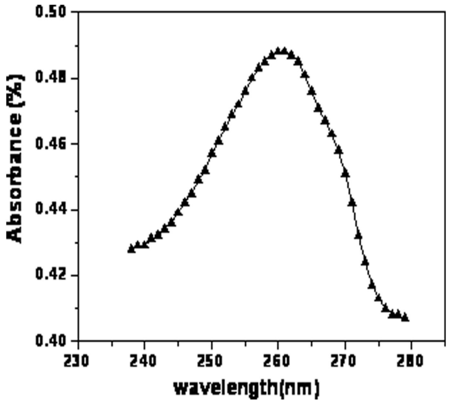


Fig. 6 – *UV of zirconium (IV) tungstate*

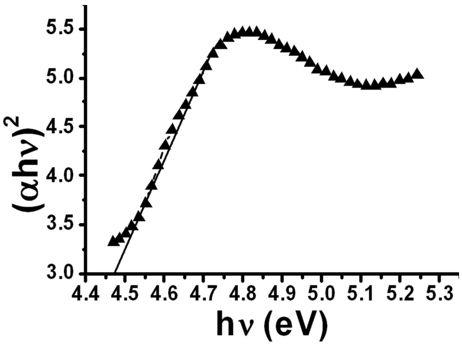


Fig. 7 – *Plot of* $(\alpha h\nu)^2$ *versus* $h\nu$

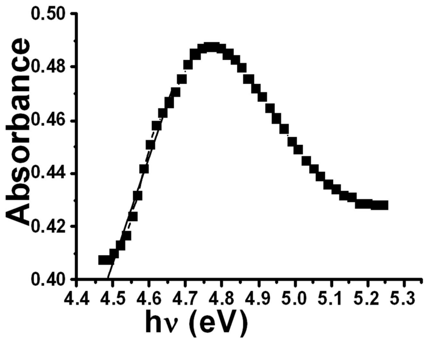


Fig. 8 – *Absorbance versus energy spectrum*

A band gap of 4.48 eV for nanocrystalline zirconium tungstate is due to ligand to metal charge transfer transition. The band gap also depends on the zirconia and tungstate content in the sample [12]. Also the energy gap is sensitive to particle size in 1 - 10 nm range for many semiconductors including zirconia and WO_3. In this size range, a large fraction of atoms residing at the surface alters the cluster electronic properties and hence the band gap value. Thus materials of desired band gap can be designed by changing the stoichiometry of Zr to W and thereby varying the particle size.

3. CONCLUSIONS

Zirconium (IV) tungstate is prepared in the nanoform by chemical coprecipitation method with high purity and homogeneity. XRD confirms the crystalline nature of the material. HRTEM and SEM reveal the spherical nature of the particles. Particle size is found to be < 50 nm from HRTEM. The energy gap determined by Tauc relation is 4.48 eV suggesting that it is a n-type semiconductor suitable for device applications which use metal oxide semiconductor such as gas sensors, UV sensors, etc.

to Contents, page 1

REFERENCES

1. Y.D. Wang, Z.X. Chen, Y.F. Li, Z.L. Zhou, and X.H. Wu, *Solid State Electron.* **45,** 639 (2001).
2. K. Huang, Q. Pan, F. Yang, S. Ni and D. He, *Physica E* **39,** 219 (2007).
3. S.M. Kanan, C.P. Tripp, *Curr. Opin. Solid St. M.* **11,** 19 (2007).
4. P.Z. Si, C.J. Choi, E. Bruck, J.C.P. Klaasse, D.Y. Geng, and Z.D. Zhang, *Physica B* **392,** 154 (2007).
5. E.B. Franke, C.L. Trimble, J.S. Hale, M. Schubert, and J.A. Woollam, *J. Appl. Phys.* **88,** 5777 (2000).
6. L.C. Chen and K.C. Ho, *Electrochim. Acta* **46,** 2151 (2001).
7. X. Liu, A. Huang, C. Ding and P.K. Chu, *Biomaterials* **27,** 3904 (2006).
8. K.N. Rao, A. Sridhar, A.F. Lee, S.J. Taveneer, N.A. Young and K. Wilson, *Green Chem.* **8,** 790 (2006).
9. I.K. Ghosh, M.K. Mitra and K.K. Chattopadhyay, *Nanotechnology* **16,** 107 (2005).
10. M.M. Bagheri-Mohagheghi, N. Shahtahmasebi, M.R. Alinejad, A. Youssefi and M. Shokooh- Saremi, *Physica B* **403,** 2431 (2008).
11. J.S.O. Evans, T.A. Mary and T. Voget, *Chem. Mater.* **8,** 2809 (1996).
12. H. Zhu, D. Yang, G. Yu, H. Zhang, and K. Yao, *Nanotechnology* **17,** 2386 (2006).
13. D.G. Barton, S.L. Soled, and E. Iglesia, *Top. Catal.* **6,** 87 (1998).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.185-192

PACS number: 73.63.Bd

EFFECT OF ANNEALING ON STRUCTURE, MORPHOLOGY, ELECTRICAL AND OPTICAL PROPERTIES OF NANOCRYSTALLINE TiO_2 THIN FILMS

S.G. Pawar, M.A. Chougule, P.R. Godse, D.M. Jundale, S.A. Pawar, B.T. Raut, V.B. Patil

Materials Research Laboratory, School of Physical Sciences,
Solapur University, Solapur-413255, M.S., India
E-mail: drvbpatil@gmail.com

Semi-transparent and highly conducting nanostructured titanium oxide thin films have been prepared by sol-gel method. Thin films of TiO_2 deposited on glass substrates using spin coating technique and the effect of annealing temperature (400 - 700 °C) on structural, microstructural, electrical and optical properties were studied. The X-ray diffraction and Atomic force microscopy measurements confirmed that the films grown by this technique have good crystalline tetragonal mixed anatase and rutile phase structure and homogeneous surface. The study also reveals that the rms value of thin film roughness increases from 7 to 19 nm. HRTEM image of TiO_2 thin film (annealed at 700 °C) shows that a grain of about 50 - 60 nm in size is really aggregate of many small crystallites of around 10 - 15 nm. Electron diffraction pattern shows that the TiO_2 films exhibited tetragonal structure. The surface morphology (SEM) of the TiO_2 film showed that the nanoparticles are fine with an average grain size of about 50 - 60 nm. The optical band gap slightly decreases from 3.26 - 3.24 eV and the dc electrical conductivity was found in the range of 10^{-6} to $10^{-5}(\Omega\ cm)^{-1}$ when the annealing temperature is changed from 400 to 700 °C. It is observed that TiO_2 thin film annealed at 700 °C after deposition provide a smooth and flat texture suited for optoelectronic applications.

***Keywords**: SOL-GEL METHOD, STRUCTURAL PROPERTIES, OPTICAL PROPERTIES, ELECTRICAL CONDUCTIVITY.*

(Received 04 February 2011, in final form 16 March 2011)

1. INTRODUCTION

Many nanostructural materials are now investigated for their potential applications in photovoltaic, electro-optical, micromechanical and sensor devices [1]. Nanoporous TiO_2 thin films for dye-sensitized and ETA (extremely thin absorber) solar cells have been under intense study for many years [2]. The TiO_2 occurs naturally as minerals rutile, anatase and brookite phases. The rutile and anatase phases have been intensively studied and have significant technological uses related in large measure to their optical properties: both are transparent in the visible range and absorb in the near ultraviolet. For the TiO_2 brookite, the width of these domains are depends on the TiO_2 thickness values [3]. The rutile (110) surface is used as a prototypical model for basic studies of oxide surfaces, and is the active component in self-cleaning cement [3]. There is a recent development of interest on transparent rutile n-doped films [4]. At room temperature, the

to Contents, page 1

direct gap energy E_g is 3.06 eV for rutile and is about 3.3 eV for anatase. Mixed-phase TiO_2 material has recently been fabricated by chemical and physical methods, including a sol-gel, hydrothermal, solvothermal, and reactive DC magnetron sputtering method, and has demonstrated excellent photocatalytic activities [5-8]. As usual, the preparation of titanium nanomaterial with two different polymorphs by sol-gel method needs to crystallize the as-prepared titanium hydroxide at high temperature ($\geq$ 700 °C). This heat treatment leads to change in the crystallite size as well as morphology of the sol-gel derived titanium oxide [9, 10].Therefore; the preparation of mixed phase titanium oxide at lower temperature can be useful both in saving energy and getting better properties [11-15]. According to the literature survey, there isn't any study which attempted to systematic investigations viz effect of annealing on structure, morphology, electrical conductivity and band gap of multiphase TiO_2 by sol-gel method.

In the present paper, we report preparation and deposition of nanocrystalline TiO_2 mixed phase (anatase and rutile) thin films by sol-gel spin coating technique. The nanopowders are subsequently sintered at 400 - 700 °C. The nanopowder and films were further investigated for their structural, microstructural and optoelectronic properties.

2. EXPERIMENTAL DETAILS

Nanocrystalline TiO_2 is synthesized by sol-gel method using titanium isopropoxide as a source of Ti. 3.7 ml of titanium isopropoxide was added to 40 ml of methanol and stirred vigorously at temperature 60 °C for **1** h, this leads to formation of white powder which was sintered at 400 - 700 °C for 1 h in air to achieve formation of nanocrystalline TiO_2 of 50 - 60 nm size. The nanocrystalline TiO_2 powder was dissolved in m – cresol. The solution was stirred for 1 h at room temperature and filtered. A thin film of this filtered nanocrystalline TiO_2 was deposited on a glass substrate by a single wafer spin processor (APEX Instruments, Kolkata, Model SCU 2007). After setting the substrate on the disk of the spin coater, the coating solution approximately 0.2 ml was dropped and spin-coated with 3000 rev. min^{-1} for 40 s in air and dried on a hot plate at 100 °C for 10 min.

The structural properties of the films were investigated by X-ray diffraction (XRD) (Philips PW – 3710, Holland) using filtered Cu K_α radiation (λ = 1.5406 Å). High resolution Transmission electron microscopy (HRTEM) and small area electron diffraction (SAED) were obtained in order to investigate the morphology and structure of titanium oxide thin films. The TEM images were taken with a Hitachi Model H-800 transmission electron microscope. In order to determine the particle size and morphology of nanopowder, the annealed powders were dispersed in m-cresol and sonicated ultrasonically by using Microclean-103 (OSCAR ultrasonic bath apparatus) to separate out individual particles. The size and morphology of the thin films were then observed on SEM Model: JEOL JSM 6360 operating at 20 kV. Roughness of the film was determined from the Atomic force microscopy (AFM) using (Digital Instruments) Nanoscope III a.

The room temperature dc conductivity measurements were made on thin films using four probe techniques. The optical absorption spectra of TiO_2 thin films were measured using a double-beam spectrophotometer Shimadzu

to Contents, page 1

UV-140 in the 200 - 1000 nm-wavelength range. The thickness of the film was measured by using profilometry using a Dektak profilometer. The values obtained ranged between 100 and 200 nm.

3. RESULTS AND DISCUSSION

3.1 Structural Properties

Fig. 1 shows the X-ray diffraction patterns of nanocrystalline TiO_2 powder annealed at 400 - 700 °C temperatures with a fixed annealing time of 1 h in air. The effect of annealing temperature on the crystallinity of TiO_2 can be understood from the figure.

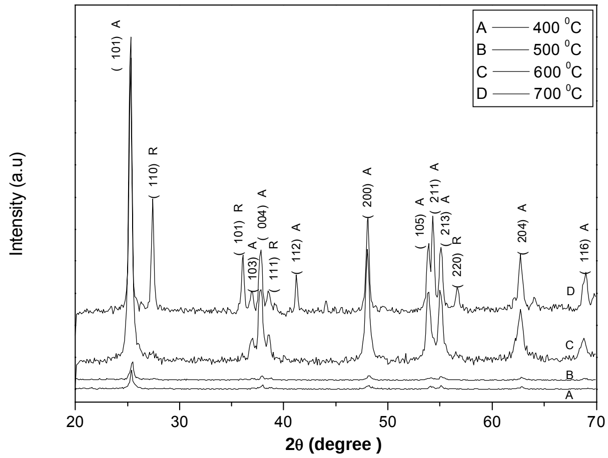


***Fig. 1** – X-ray diffraction patterns of TiO_2 nanopowder at different annealing temperatures: (a) 400 °C (b) 500 °C, (c) 600 °C and (d) 700 °C*

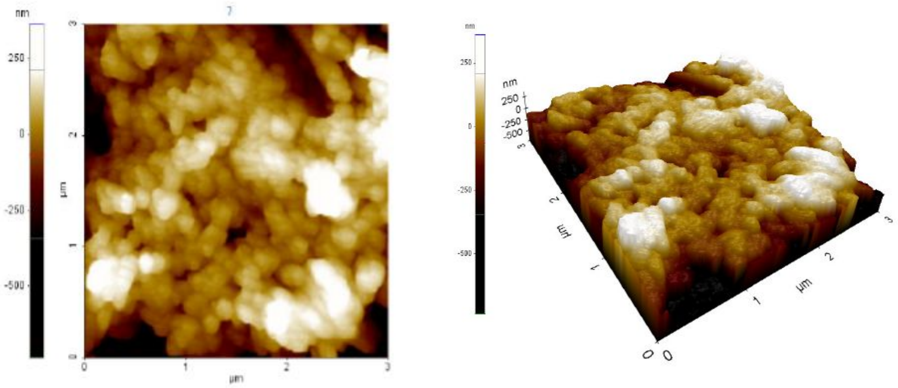
The X-ray spectra show well-defined diffraction peaks showing good crystallinity. The crystallites are randomly oriented and the d-values calculated for the diffraction peaks are in good agreement with those given in JCPD data card (JCPDS No. 78 – 2485 & 78 – 2486) for TiO_2 anatase and rutile. This means that TiO_2 has been crystallized in a tetragonal mixed anatase and rutile form. These results are in good agreement with other reports on the mixed phase TiO_2 by sol gel method [5-8]. The lattice constants calculated from the present data are $a = 3.7837$ Å and $c = 9.5087$ Å respectively. From Fig. 1 it is seen that the (101) peak (anatase phase) of intensity increased with an increase in the annealing temperature. However, the full width at half-maxima FWHM of the (101) peaks was hardly changed with increasing film annealing temperature [9, 10]. The grain size of all TiO_2 samples sintered at 400 °C to 700 °C was calculated using Scherer's equation and it is in the range of 50 - 60 nm, revealing a fine nanocrystalline grain structure.

3.2 Microstructural and morphological properties

AFM (non contact mode) was used to record the topography of the nanocrystalline TiO_2. In this mode, the tip of the cantilever does not contact

to Contents, page 1

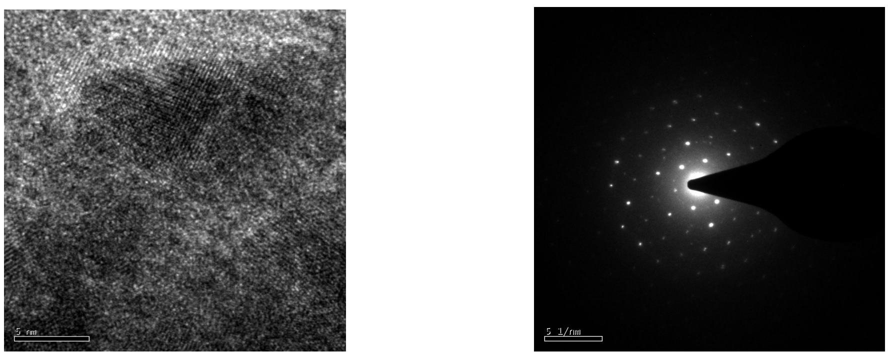
the sample surface. The cantilever is instead oscillated at a frequency slightly above its resonance frequency where the amplitude of oscillation is typically a few nanometers (< 10 nm). The surface morphologies of the TiO_2 nanoparticles exhibit notable features. Figures 2 show 2D and 3D AFM images (3 μm $\times$ 3 μm) of the TiO_2 nanoparticle films. The average surface roughnesses are in 7 - 19 nm for TiO_2 nanoparticles. The average particle sizes of TiO_2 are found to be in the range of 50 - 60 nm.



a b

Fig. 2 – *AFM images of TiO_2 thin films annealed at 700 °C for 1 h in air (a) planer view (b) 3D view*

Figure 3 a show high resolution image of titanium oxide thin film annealed at 700 °C, recorded from typical regions of films. HRTEM shows a large number of crystalline grains appear in a structured matrix and the grains have a diameter in the range 2 - 3 nm and show lattice spacings of about 0.35 nm. Figure 3 b shows the diffraction patterns obtained from the titanium oxide film annealed at 700 °C. The different arrangement of dominant diffracted rings indicates a phase evolution of crystalline grains because of thermal annealing.



a b

Fig. 3 – *TEM of TiO_2 thin film annealed at 700 °C for 1 h in air (a) Microstructure (b) Selected Area Electron Diffraction pattern*

to Contents, page 1

Table 1 shows the interplanar spacings determined from diffraction pattern together with the corresponding ones of TiO_2 rutile and anatase phases reported in the literature for comparison [9, 10]. It is evident that the film consists of grains having both the interplanar spacings of TiO_2 in the anatase and rutile structural modifications [5-8].

Table 1 *– Interplanar spacings deduced from electron diffraction patterns reported in Fig. 3 b together with the corresponding ones obtained from literature data. Numbers in brackets (n) represent the labels of reflections in diffraction patterns.*

Interplanar spacings determined titanium oxide films		Interplanar spacings in this work on reported in literature	
(n)	*d* (nm)	Rutile *d* (nm) – hkl	Anatase *d* (nm) – hkl
(1)	0.3551		0.35126 - 101
	0.3242	0.3246 - 110	
		0.25130 - 101	
(2)	0.2371		0.23775 - 004
		0.2187 - 111	
(3)	0.1894		0.18900 - 200
(4)	0.1665		0.16690 - 105
(5)	0.1661		0.16643 - 211
			0.14754 - 204

[a]Joint commission Powder Diffraction File No. (78) – 2485 & 78) – 2486)

The film microstructure was studied using Scanning Electron Microscopy. Fig. 4 show the SEM morphology of TiO_2 thin film annealed at different temperatures 400 - 700 °C.

In general, films are homogeneous and continuous separate coating layers are not visible in annealed films. There seems to be mismatch in average size of grains / particles determined through Scherer's calculation utilizing XRD data and SEM analysis .SEM image suggest size of grains to be much larger. Further, while Scherer's calculation suggests an increase in particle size with rise in annealing temperature, SEM images indicate almost a reverse trend. Taking into account the above discrepancy and the fact that SEM analysis reveals formation of particles with different shapes and size, it seems appropriate to consider that the particles which appears in SEM images are, in fact, grain agglomerates, which get fragmented with rise in annealing temperature. In films annealed at 400, 500, and 600 °C, relatively larger particles / grain agglomerates can be seen compared to films annealed at 700 °C, the film morphology appeared most uniform and the particle size also lowest [9, 10].

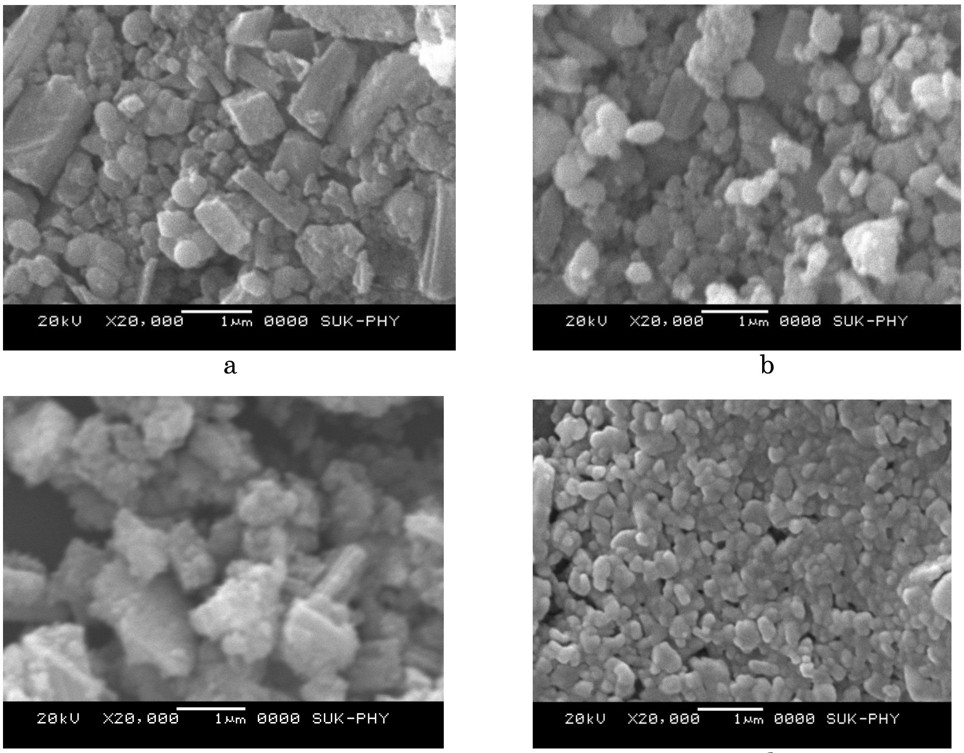


***Fig. 4** – SEM image of* TiO_2 *thin film at different annealing temperatures: (a)* 400 *°C, (b)* 500 *°C, (c)* 600 *°C, and (d)* 700 *°C*

3.3 Electrical Properties

The dc electrical conductivities σ_{dc} of TiO_2 films annealed at 400 - 700 °C as a function of reciprocal temperature were measured in the 300 - 600 K temperature range and their temperature dependence can be fitted to a usual Arrhenius equation:

$$\sigma_{dc} = \sigma_0 \exp(-E_{a\sigma}/k_B T) , \tag{1}$$

where $E_{a\sigma}$ is the conductivity activation energy and is k_B Boltzmann constant. The temperature dependence of electrical conductivity (Fig. 5) showed two distinct conduction regions corresponding to two different conduction mechanisms; one, a grain boundary scattering limited and second a variable range hopping [16]. The activation energies of an electrical conduction have been computed for both regions and their variation with annealing temperature is shown in Table 2.

to Contents, page 1

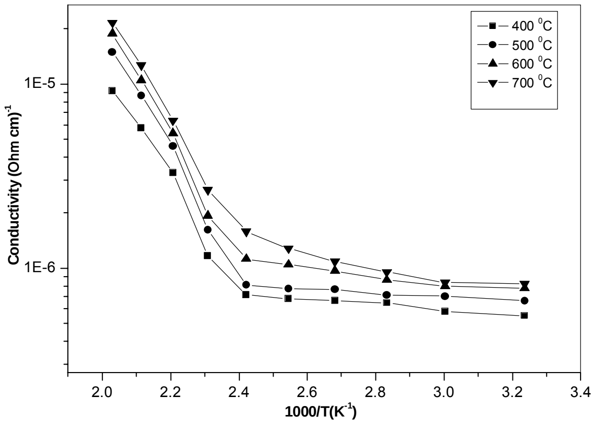


Fig. 5 – *Arrhenius plot of dc conductivity vs.* 1000/*T of* TiO_2 *thin film annealed at (a)* 400 °C *(b)* 500 °C, *(c)* 600 °C *and (d)* 700 °C *for 1 h in air*

Table 2 – *Effect of annealing on thin film properties of* TiO_2

№	Annealing temperature, °C	Crystallite Size, nm	Energy gap E_g, eV	Activation energy, E_a eV	
				LT	HT
1	400	50	3.26	0.746	0.195
2	500	54	3.25	0.730	0.202
3	600	59	3.24	0.703	0.180
4	700	60	3.24	0.696	0.190

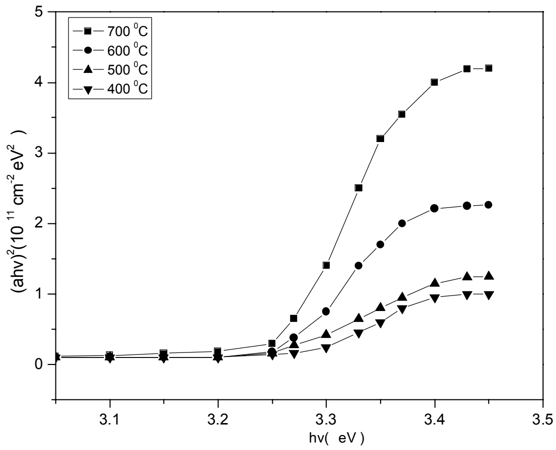


Fig. 6 – *Plot of* $(\alpha h\nu)^2$ *versus* $(h\nu)$ *for different annealing temperatures*

3.4 Optical Properties

The optical constants namely absorption coefficient (σ), energy gap (E_g) and the type of the optical transition have been determined by examining an optical absorption spectrum in the 200 - 1000 nm wavelength range at different annealing temperature. Fig. 6 shows determination of the optical

to Contents, page 1

gap from the $(\alpha h\nu)^2$ versus $(h\nu)$ variation in the linear region. The values of optical band gap decreasing slightly with increasing annealing temperature 400 - 700 °C (Table. 2). It is found that the optical absorption coefficient is larger for all the films ($\approx 10^4$ cm^{-1}). This may be accounted for the fact that the quality of the TiO_2 film improves when the sample is annealed at a higher temperature (in this case 700 °C). This fact is also by the XRD, SEM images shown in Figs. 1, 4a-d. The $(\alpha h\nu)^2$ versus hυ plots show straight line behaviour on the higher energy side shows direct type of transitions involved in these films. The type of transitions for TiO_2 films annealed at 400 - 700 °C was confirmed by plotting $\ln(\alpha h\nu)$ versus $\ln(h\nu - E_g)$ variation [17-18].

4. CONCLUSIONS

Nanocrystalline titanium oxide thin films were prepared by sol-gel spin coating techniques on the glass substrate. The effect of annealing temperature on structure, microstructure, morphology ,electrical and optical properties of TiO_2 thin films were studied by XRD, HRTEM, AFM, SEM, four probe and UV-Visible measurements .The XRD results reveal that the deposited thin film of TiO_2 has a good nanocrystalline tetragonal mixed anatase and rutile phase structure. The HRTEM, AFM and SEM results demonstrate that a uniform surface morphology and the nanoparticles are fine with an average grain size of about 50 - 60 nm. The dc electrical conductivity found in the range of 10^{-5} to 10^{-6} $(\Omega\ \text{cm})^{-1}$. Optical studies showed that the TiO_2 has high absorption coefficient ($\approx 10^4$ cm^{-1}) with a direct band gap. The optical band gap decreasing slightly with increasing annealing temperature (3.26 - 3.24 eV).

The author (VBP) thankful to Department of Science and Technology, Government of India, New Delhi for sanctioning Fast Track Project (SR/FTP/PS-09/2007).

REFERENCES

1. R.S. Singh, V.K. Rangari, S. Sanagapalli, V. Jayaraman, S. Mahendra, V.P. Singh, *Sol. Energ. Mat. Sol. C.* **82,** 315 (2004).
2. J. Sabataityte, I. Oja, F. Lenzmann, O. Volobujeva, M. Krunks, *CR Chim.* **9,** 708 (2006).
3. S.S. Madaeni, N. Ghaemi, *J Membrane Sci.* **303,** 221 (2007).
4. A.R. Gandhe, J.B. Fernandes, *J Solid. State Chem.* **178,** 2953 (2005).
5. K.B. Sundaram, A. Khan, *Thin Solid Films* **295,** 87 (1997).
6. J. Aranovich, A. Ortiz, R.H. Bube, *J. Vac. Sci. Technol.* **16,** 994 (1979).
7. M.N. Islam, M.O. Hakim, H. Rahman, *J. Mater. Sci.* **22,** 1379 (1987).
8. S. Major, K.L. Chopra, *Solar Energ. Mater.* **17,** 319 (1988).
9. M. de la L. Olvera, A. Maldonado, R. Asomoza, M. Konagai, M. Asomoza, *Thin Solid Films* **229,** 196 (1993).
10. K. Kamata, S. Matsumoto, P. Souleite, B.W. Wessels, *J. Mater. Res.* **3,** 740 (1988).
11. Y. Natsume, H. Sakata, T. Hirayama, H. Yanagida, *J. Appl. Phys.* **72,** 4203 (1992).
12. W. Tang, D.C. Cameron, *Thin Solid Films* **238,** 83 (1994).
13. V. Craciun, J. Elders, J.G.E. Gardeniers, I.W. Boyd, *Appl. Phys. Lett.* **65,** 2963 (1994).
14. M.M.A. Sekkina, A. Tawfik, M.I. Abd El-Ati *Thermochimica Acta,* **86,** 59 (1985).
15. S. Chandra, R.K Pandya, *Phys. Status Solidi A* **72,** 415 (1982).
16. K.L. Narasimhan, S.P. Pai, V.R. Palkar, R. Pinto, *Thin Solid Films* **295,** 104 (1997).
17. A.P. Roth, D.F. Williams, *J. Appl. Phys.* **52,** 6685 (1981).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.193-202

PACS numbers: 61.05.C –, 68.37.Lp

STUDY ON NANOPARTICLES OF ZnSe SYNTHESIZED BY CHEMICAL METHOD AND THEIR CHARACTERIZATION

M.P. Deshpande, S.H. Chaki, N.H. Patel, S.V. Bhatt, B.H. Soni

Department of Physics,
Sardar Patel University, 388120, Vallabh Vidyanagar, Gujarat, India
E-mail: vishwadeshpande@yahoo.co.in

The properties of semiconductor nanoparticles depend mainly on their shape and size due to high surface-to-volume ratio. The II – VI semiconductors have many applications such as, LED, acousto-optical effects and biological sensors. The ZnSe nanoparticles have wide-ranging applications in laser, optical instruments etc. because it has wide band gap and transmittance range, high luminescence efficiency, low absorption coefficient. In recent years, much attention was paid on the preparation methods, performances and applications of ZnSe nanoparticles and thin solid films, and a lot of important accomplishments have been obtained. In the present study ZnSe nanoparticles were successfully prepared by reacting $Zn(CH_3COO)_2{\cdot}2H_2O$ and Na_2SeSO_3 at 343 K. The size of the crystallite was estimated by X-ray diffraction and TEM, whereas EDAX has confirmed of no foreign impurity inclusion in ZnSe nanoparticles. XRD shows the crystallite size of 5.68 nm and TEM gives a distribution ranging from 20 nm to 71 nm. A SEM image shows that the particles are spherical in a shape. Quantum confinement has resulted in the blue shift compared to bulk ZnSe as observed from the absorption spectra of particles dispersed in DMF. We obtained the photoluminescence spectra on these particles with two different excitation wavelength which shows broad band emission peak at 573 nm. Photoluminescence spectra taken with other excitation wavelength also gives sharp emission peaks at 484 nm, 530 nm, 551 nm and 600 nm.

Keywords: *ZnSe NANOPARTICLES, CHEMICAL METHOD.*

(Received 04 February 2011, in final form 18 March 2011)

1. INTRODUCTION

Several approaches and a lot of publications have been devoted to prepare and control the size and shape of II – VI semiconductor nanoparticles. Among all semiconductor materials, the II – VI semiconductor systems have many applications, such as light emitting diodes and acousto-optical effects and biological sensors [1-5]. Zinc selenide (ZnSe) is an important II – VI semiconductor material, with a rather wide bulk direct bandgap energy (E_g = 2.7 eV) [6-7], and a small exciton Bohr diameter of 9 nm [12]. Zinc selenide is normally known to be an n-type semiconducting material [6-11].The synthesis of nanocrystalline ZnSe powders with tunable phase, morphology and size provides alternative variables in tailoring its physical and chemical properties [13-17]. The ZnSe nanoparticles have wide-ranging applications in laser, optical instruments, etc. because it has wide band gap (2.58 eV) and transmittance range (0.5 - 22 μm), high luminescence efficiency, low absorption coefficient, and excellent transparency to infrared

to Contents, page 1

[18-20]. It is used to increase the open circuit voltage of solar cells [21]. It exhibits great potential for various optoelectronic and high-speed applications like blue-green light emitting diodes, photoluminescent and electroluminescent devices, lasers and thin film solar cells [22-23].

ZnSe nanocrystalline samples were mostly prepared by molecular beam epitaxy (MBE) [24], metalorganic chemical vapor deposition (MOCVD) [25] and organometallic vapor phase epitaxy (OMVPE) [26], Solvothermal or hydrothermal route [27-28], mechanochemical synthesis from Zn and Se granules [29]. Now days, a wide variety of methods have been employed for preparing zinc selenide (ZnSe) nanocrystals or nanoparticles with controlled size such as solvothermal [30], sol-gel method [31], and microemulsion technique [32]. An effective control over the particle size can be achieved by manipulating different parameters namely, concentration, molar ratio of the reactants, time of heating in chemical method [33, 34] and radiation doses in radiolytic method [35-36]. Recently, a variety of ZnSe nanostructured materials in various geometrical morphologies, including nanoparticles [37], rods [38], wires [39], belts [40], needles [41], saw and tube [42], plates [43], hollow microspheres constructed with nanoparticles [44], and flowerlike pattern of radially aligned nanoflakes [45], nanoribbons[46], nanowheels [47], nanorings [48], nanopowders [49] have successfully been prepared by various methods.

2. EXPERIMENTAL DETAIL

Chemicals used for the synthesis of ZnSe nanoparticles were zinc acetate [$Zn(CH_3COO)_2$], sodium selenosulphate (Na_2SeSO_3), ammonia solution (25 %) [NH_3], hydrazine hydrate (80 %) [$N_2H_4.H_2O$], sodium hydroxide [NaOH], selenium powder (99 % purity) [Se], sodium sulfite [Na_2SO_3]. All chemicals used were A.R. grade. Here zinc acetate work as Zn^{2+} ion source, Na_2SeSO_3 as Se^{2-} ion source, ammonia as a complexing agent, hydrazine hydrate as catalyst and sodium hydroxide as a pH adjustor.

The 0.25 M sodium selenosulphate solution was prepared by mixing 2.36 gm selenium powder (99% purity) with 9.48 gm anhydrous sodium sulfite in 120 ml of distilled water with constant stirring for 8 h. It was sealed and kept overnight, since on cooling, a little selenium separated out from the solution. It was then filtered to obtain a clear solution. A 25 ml of (0.5 M) zinc acetate solution was taken in a beaker of 200 ml capacity. To, it 1ml of hydrazine hydrate (80 %) was added under a constant stirring. To this, under constant stirring, a sufficient amount of ammonium solution (25 %) was added to dissolve the turbidity of resultant solution. The desired pH of the resultant solution was adjusted by adding sodium hydroxide (1 M) solution. Then the reactant vessel was kept at 343 K in a constant temperature water bath. When appropriate temperature of 333 K was reached, sodium selenosulphate (0.25 M, 25 ml) solution was added to the bath. Then heated continuously at constant 343 K temperature for 3 - 4 hour and cooled to the room temperature. White colored precipitates obtained at the bottom of the beaker were filtered, washed a number of times in distilled water and dried in a vacuum.

to Contents, page 1

3. RESULT AND DISCUSSION

3.1 Chemical mechanism

Sodium selenosulphate hydrolysis in the solution to give Se^{2-} ions according to,

$$Na_2SeSO_3 + OH^- \rightarrow Na_2SO_4 + HSe \quad (1)$$

$$HSe^- + OH^- \rightarrow H_2O + Se^{2-}. \quad (2)$$

When hydrazine hydrate is added in Zinc Acetate it forms the ions of Zn^{+2}

$$mZn(CH_3COO)_2 + nN_2H_4 \rightarrow [Zn(CH_3COO)_2]_m \cdot [N_2H_4]_n \quad (3)$$

$$[Zn(CH_3COO)_2]_m \cdot [N_2H_4]_n \rightarrow mZn^{2+} + nN_2H_4 + 2m(CH_3COO)^{2-}. \quad (4)$$

When ammonia solution is added to this Zn-salt solution it gives the complex of zinc tetra-amine ion $[Zn\,(NH_3)_4{}^{2+}]$ as,

$$Zn^{2+} + 4NH_3 \rightarrow Zn\,(NH_3)_4^{2+}. \quad (5)$$

Then $Zn\,(NH_3)_4{}^{2+}$ react with Se^{2-} ions that result in the formation of ZnSe nanoparticles as follows:

$$Zn \cdot (NH_3)_4^{2+} + Se^{2-} \rightarrow ZnSe + 4NH_3. \quad (6)$$

3.2 Energy dispersive analysis of X-ray (EDAX)

The compositions of synthesized ZnSe nanoparticles were determined by the spectra obtained by Energy dispersive analysis of X-rays (EDAX) which is shown in Fig. 1. This analysis was carried out using Philips EM 400 electron microscope. It is reflected from the EDAX spectra that no foreign impurities are present in the sample.

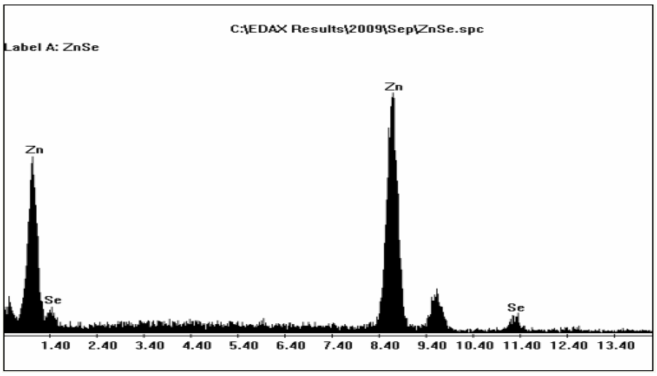


***Fig. 1** – EDAX Spectra of ZnSe nanoparticles*

3.3 X-ray diffraction (XRD)

The X-ray diffractogram of ZnSe nanocrystallites synthesized by chemical method is shown in Fig. 2 and indexed based on hexagonal system. This analysis was carried out using XRD diffractometer (powder) Philips Xpert MPD. The values of the lattice parameter determined from the X-ray diffractogram using powder X software which clearly matches with the reported values of lattice parameter of

to Contents, page 1

ZnSe (JCPDS No. 15-0105). The presence of a relatively sharp peak on a background of a wider peak suggest that both large and small grains are collected together. The peaks (N) are not identified as ZnSe and may be due to zinc complex formed during the reaction [50].

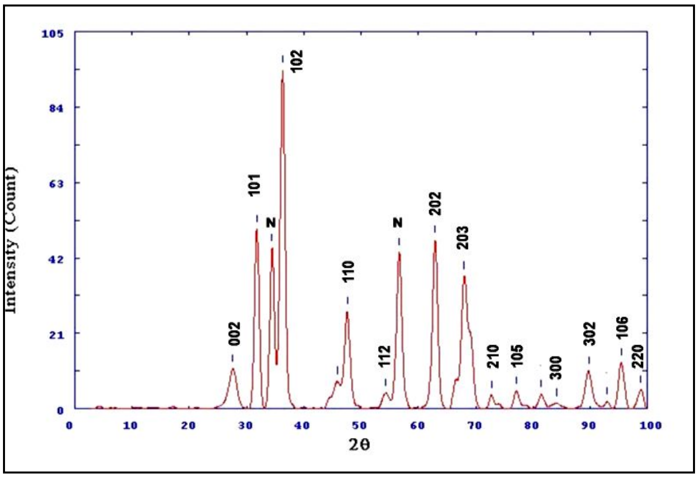


Fig. 2 *– X-ray diffractogram of ZnSe nanocrystallites*

We used Scherrer's formula [51-53] for calculating the crystallite size which is given by,

$$t = k\lambda/\beta_{2\theta}\cos\theta$$

where, t is the crystallite size, k is the Scherrer constant whose value is chosen to be unity by assuming the particle to be spherical, λ is the wavelength of X-ray beam, θ is the Bragg angle. $\beta_{2\theta}$ is the width at half the maximum intensity measured in radians.

$$\beta_{2\theta} = \sqrt{(FWHM_{obs})^2 - (FWHM_r)^2}$$

$FWHM_{obs}$ is observed reflection from ZnSe nanocrystallite. $FWHM_r$ is instrumental function of the diffractometer [54].

This technique gives us the rough estimate of crystallite size ranging from 5 nm to 9 nm due to correction factor involved in the Scherrer's equation. Fig. 3 shows Hall-Williamson plots for ZnSe nanocrystallites. The most widespread method used for estimating the nanocrystallite sizes is the calculation of line broadening by the Williamson-Hall method [55], which not only uses the Scherrer's formula [51-53] but also takes into account the microstrain of nanocrystals. This plot requires large number of nonoverlapping reflections. For this reason, we used the Williamson-Hall method to calculate the size of particles and also to obtain information of the strain from the full width at half maximum (FWHMs) of the diffraction peaks. The FWHMs (β) can be expressed as a linear combination of the contributions from the strain (ε) and crystallite size (L) through the following relation [56].

to Contents, page 1

$$\frac{\beta\cos\theta}{\lambda}=\frac{K}{L}+\frac{4\varepsilon\sin\theta}{\lambda}$$

The plot of $\beta{\cdot}\cos\theta/\lambda$ versus $(\sin\theta/\lambda)$ for ZnSe nanocrystallites which is straight line. The slope of the plot gives the amount of residual strain, which turns out to be − 0.0026 for ZnSe nanocrystallite. The reciprocal of intercept on the $(\beta{\cdot}\cos\theta/\lambda)$ axis gives the average crystallite size as ~ 5.68 nm. The negative value of residual strain indicates the presence of compressive strain in ZnSe nanocrystallites.

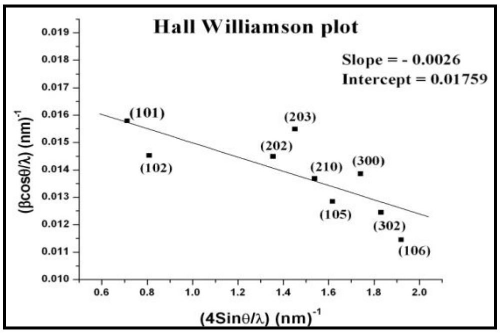
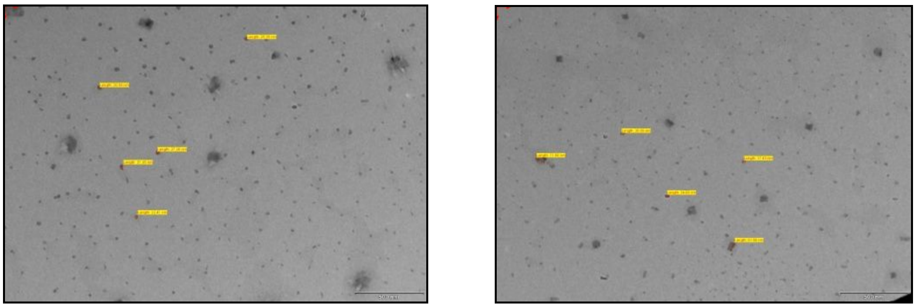


Fig. 3 *– Hall-Williamson plot for ZnSe nanocrystallite*

3.4 Transmission electron microscopy (TEM)

The TEM image of ZnSe nanoparticles is shown in Fig. 4. This analysis was carried out using Philips, Tecnai 20 microscope. The TEM morphology shows particles have spherical shape. Nanoparticle sizes determined from Fig. 4a ranges between 20 - 31 nm and Fig. 4b ranges between 17 - 71 nm which are showing higher values than those obtained from Scherrer's equation used in X-ray diffractogram. This indicates that the size distribution of nanocrystallites is not uniform and varies from 17 nm to 71 nm.



a b

Fig. 4 *– Shows images of ZnSe nanoparticles (a) and (b)*

Fig. 5 shows diffraction pattern which is indicating that ZnSe nanocrystallites are polycrystalline in nature. This diffraction pattern shows (112), (102), (202), (210) reflections which are corresponding to the hexagonal phase of ZnSe nanocrystallites and there by matching with the X-ray diffraction pattern.

to Contents, page 1

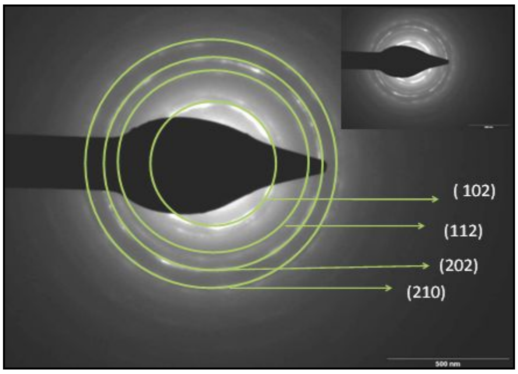


Fig. 5 *– Selected area diffraction pattern for ZnSe nanocystallites*

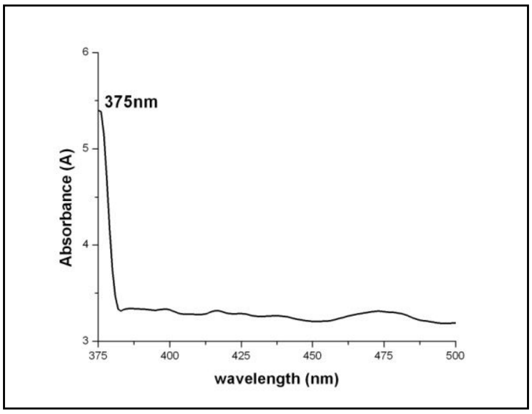


Fig. 6 *– Absorption spectra for ZnSe nanoparticles*

3.5 Absorption spectra of ZnSe nanoparticles

Absorption spectroscopy is a good way to study any change in the particle size which could occur during the aggregation of the particles. The ZnSe nanoparticles synthesized were dispersed in DiMethylFormamide (DMF) to obtain the absorption spectra in the UV-VIS region. This analysis was carried out using Perkin-Elmer Lambda 19 spectrophotometer. Fig. 6 shows the absorption spectra for ZnSe nanoparticles and it can be seen that the absorption edge lies at 375 nm with the band gap at 3.3 eV. Comparing with that of bulk material (460 nm, 2.7 eV), it is believed that the blue shift in the absorption peak was caused by the quantum confinement effect.

The size of the prepared particles can be calculated from the spectral data using Wang equation [53] given as,

$$\Delta E = \left(\frac{\hbar^2\pi^2}{2R^2}\right)\left(\frac{1}{m_e}+\frac{1}{m_h}\right)-\frac{1.8e^2}{pR}$$

which reduces as,

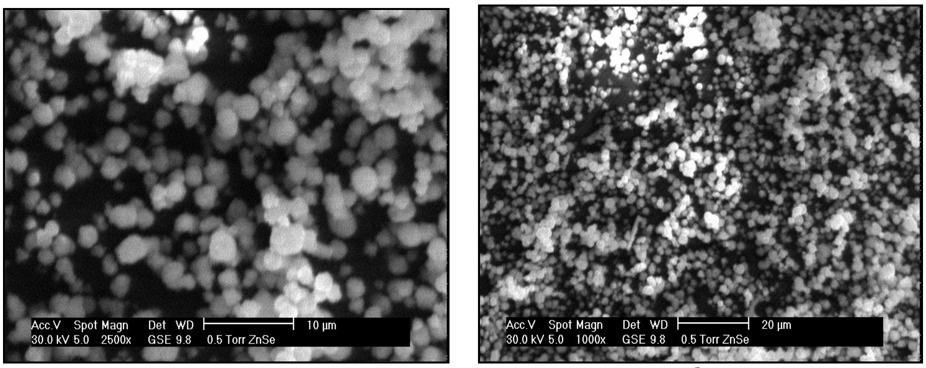
to Contents, page 1

$$E = \sqrt{E_g^2 + 2E_g h^2 \left(\frac{\pi^2}{R^2 m^*} \right)},$$

where, E_g is the bulk band gap, E is the energy gap of the size quantized ZnSe, 'R' is the radius of the particle and m^* is the effective mass with $m_e = 0.21\ m_o$ and $m_h = 0.06\ m_0$ for ZnSe [61]. By substituting our values of $E_g = 2.7$ eV for bulk ZnSe [59, 60] and $E = 3.3$ eV for ZnSe nanoparticles obtained from absorption spectra, we get the diameter of particles as 11.97 nm which shows close agreement with the values obtained from the XRD data but not with TEM.

3.6 Scanning electron microscope

The Scanning electron microscopy (SEM) image of the ZnSe nanoparticles is shown in Fig. 7a and Fig. 7b. These figures show that the product particles are spherical in shape and agglomeration is also observed. This agglo-meration is caused because we did not use any capping agent for synthesis of ZnSe nanoparticles.



a b

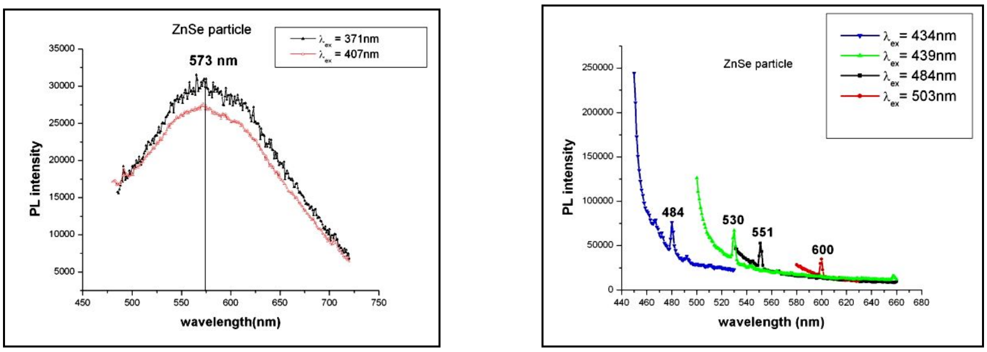
Fig. 7 *– SEM micrograph of ZnSe nanoparticles*

3.7 Photoluminescence of ZnSe nanoparticles

Photoluminescence spectra for ZnSe nanoparticles is shown in Fig. 8a. This analysis was carried out using FluoroMax – 4 Spectrofluorometer. It shows broad band emission between 500 to 650 nm which is due to deep level emission [62]. Photoluminescence spectra of ZnSe nanoparticles taken by us with two different excitation wavelength ($\lambda_e = 371$, 407 nm) is shown in Fig. 8a. It shows broad band emission peak at 573 nm with a red shift about 94 and 58 nm compared to bulk ZnSe at 465 nm respectively [63-65].This emission at 573 nm exhibits a Stokes shift.

We also took photoluminescence spectra of ZnSe nanoparticles with other excitation wavelength also ($\lambda_e = 434$ nm, 439 nm, 484 nm, 503 nm) as shown in Fig. 8b. Here we get sharp emission peak at 484 nm, 530 nm, 551 nm, 600 nm when it excited at 434 nm, 439 nm, 484 nm, 503 nm respectively and red shift of about 19 nm, 65 nm, 86 nm, 135 nm is observed compared to bulk ZnSe at 465 nm [63-65].

to Contents, page 1



a b

***Fig. 8** – (a) and (b) shows PL spectra of ZnSe nanoparticles*

4. CONCULSION

EDAX analysis confirms that the synthesized nanoparticles of ZnSe and do not contain any foreign element in them. Powder X-ray diffractogram shows that ZnSe nanocrystallites are polycrystalline in nature and belong to the hexagonal phase. The crystallite size calculated by using Scherrer's formula comes out to be between 5 - 9 nm and the lattice parameter is about a = 3.98 Å and c = 6.55 Å° which matches nearly with JCPDF files. From the Hall-Williamson plot the value of strain comes out to be – 0.0026 and average crystallite size is 5.68 nm. The TEM image shows that the particle size ranges between 17 - 71 nm and selected area diffraction pattern indicated that the synthesized ZnSe nanocrystallites are polycrystalline in nature. The absorption spectra show blue shift in the absorption edge at 375 nm for ZnSe nanoparticles in comparison to bulk ZnSe which is observed at 460 nm. The SEM image shows that the ZnSe nanoparticles are spherical in shape with agglomeration The Photoluminescence spectra shows red shift at 573 nm for ZnSe nanoparticles in comparison to bulk ZnSe which is observed at 465 nm. Another Photoluminescence spectra show red shift at 484 nm, 530 nm, 551 nm, 600 nm for ZnSe nanoparticles in comparison to bulk ZnSe which is observed at 465 nm.

ACKNOWLEGMENT

All the authors are thankful to the Sophisticated Instrumentation Centre for Applied Research & Testing (SICART), Vallabh Vidyanagar, Gujarat for EDAX, XRD, TEM UV-VIS -NIR Spectrophotometer and SEM analysis. The authors are also grateful to the Faculty of Engineering and Technology, M. S.University of Baroda, Vadodara for photoluminescence analysis on our samples.

REFERENCES

1. V.L. Colvin, M.C. Schlamp, A.P. Alivisatos, *Nature* **370,** 354 (1994).
2. W.C.W. Chan, S. Nie, *Science* **281,** 2016 (1998).
3. I.V. Kityk, J. Kasperczyk, B. Sahraoui, M.F. Yasinskii, *Polymers* **38,** 4803 (1997).
4. G.P. Mitchell, C.A. Mrkin, R.L. Letsinger, *J. Am. Chem. Soc.* **121,** 8122 (1999).
5. I.V. Kityk, Phys. Solid State **33,** 1026 (1991).
6. M.A. Russak, *J. Electrochem. Soc.* **129,** 542 (1982).

to Contents, page 1

7. O.V. Krylov, *Catalysis by Nonmetals* (Academic Press: New York: 1970).
8. J. Gautron, P. Lemmason, C. Rabuyo, R. Tribaulet, *J. Electrochem. Soc*. **126,** 1868 (1979).
9. P. Lemasson and J. Gautron, *J. Electroanal. Chem*. **119,** 289 (1989).
10. P. Lemasson, *J. Cryst. Growth* **72,** 405 (1985).
11. G. Hodes, D. Cahen, J. Manassen, M. David, *J. Electrochem. Soc*. **127,** 544 (1980).
12. N. Kumbhojkar, S. Mahamuni, V. Leppert. S.H. Risbud, *Nanostruct. Mater*. **10** No2, 117 (1998).
13. A.B. Panda, S. Acharya, S. Efrima, Y. Golan, *Langmuir* **23,** 765 (2007).
14. H. Gong, H. Huang, L. Ding, M.Q. Wang, K.P. Liu, *J. Cryst. Growth* **288,** 96 (2006).
15. Y. Jiao, D.B. Yu, Z.R. Wang, K. Tang, X.Q. Sun, *Mater. Lett*. **61,** 1541 (2007).
16. H.N. Wang, F.L. Du, *Cryst. Res. Technol*. **41,** 323 (2006).
17. Q. Peng, Y.J. Dong, Y.D. Li, *Angew. Chem. Int. Ed*. **115,** 3135 (2003).
18. A.P. Alivisatos, *Science* **271,** 933 (1996).
19. H. Luo, J.K. Furdyna, *Semicond. Sci. Technol*. **10,** 1041 (1995).
20. Q. Peng, Y.J. Dong, Y.D. Li, *Angew. Chem. Int. Ed*. **42,** 3027 (2003).
21. T.L. Chu, S.S. Chu. *Prog. Photovoltaics Res. Appl*. **1,** 31 (1993).
22. S.T. Lakshmikumar, A.C. Rastogi, *Thin Solid Films* **259,** 150 (1995).
23. C.D. Lokhande, P.S. Patil, H. Tributsch, A. Ennaoni, *Sol. Energ. Mat. Sol. C.* **55,** 379 (1998).
24. B.P. Zhang, T. Yasuda, Y. Segawa, H. Yaguchi, K. Onabe, E. Edamatsu, T. Itoh, *Appl. Phys. Lett*. **70,** 2413 (1997).
25. C.E.D. Bourret, *Appl. Phys. Lett*. **68,** 2418 (1996).
26. M.C. Klein, F. Hache, D. Ricard, C. Flytzanis, *Phys. Rev. B* **42,** 11123 (1990).
27. Y.D. Li, Y. Ding, Y.T. Qian, Y. Zhang, L. Yang, *Inorg. Chem*. **37,** 2844 (1998).
28. Z.X. Deng, C. Wang, X.M. Sun, Y.D. Li, *Inorg. Chem*. **41** 869 (2002).
29. M. Abdel Rafea, *J. Sci.-Mater. El*. **18,** 415 (2007).
30. H. Gong, H. Huang, L. Ding, M. Wang, K. Liu, *J. Cryst. Growth* **288,** 96 (2006).
31. L.S. Li, N. Pradhan, Y. Wang, X. Peng, *Nano Lett*. **4,** 2261 (2004).
32. G.N. Karanikolos, P. Alexandridis, G. Itskos, A. Petrou, T.J. Mountziaris, *Langmuir* **20,** 550 (2004).
33. D.Crouch, S. Norager, P. O'Brien, J.H. Park, N. Pickett, *Phil. Trans. R. Soc. Lond. A* **361,** 297 (2003).
34. A. Priyam, A. Chatterjee, S.C. Bhattacharyya, A. Saha, *J. Cryst. Growth* **304,** 416 (2007).
35. A. Datta, A. Priyam, S. Chatterjee, A.K. Sinha, S.N. Bhattacharya, A. Saha, *Colloids Surf. A* **301,** 239 (2007).
36. A.H. Souici, N. Keghouche, J.A. Delaire, H. Remita, M. Mostafavi, *Chem. Phys. Lett*. **422,** 25 (2006).
37. W.Z. Wang, Y. Geng, P. Yan, F.Y. Liu, Y. Xie, Y.T. Qian, *J. Am. Chem. Soc.* **121,** 4062 (1999).
38. J. Yang, C. Xue, S.H. Yu, J.H. Zeng, Y.T. Qian, *Angew. Chem.* **41,** 4697 (2002).
39. X.T. Zhang, Z. Liu, Y.P. Leung, Q. Li, S.K. Hark, *Appl. Phys. Lett.* **83,** 5533 (2003).
40. T.Y. Zhai, H.Z. Zhong, Z.J. Gu, A.D. Peng, H.B. Fu, Y. Ma, Y.F. Li, J.N. Yao, *J. Phys. Chem. C* **111,** 2980 (2007) .
41. H.Z. Fu, H.Y. Li,W.Q. Jie, L. Yang, *J. Cryst. Growth* **289,** 440 (2006).
42. A. Colli, S. Hofmann, A.C. Ferrari, C. Ducati, F. Martelli, S. Rubini, A. Franciosi, J. Robertson, *Appl. Phys. Lett*. **86,** 153103 (2005).
43. Z.X. Deng, C. Wang, X.M. Sun, Y.D. Li, *Inorg. Chem*. **41,** 869 (2002).
44. Q. Peng, Y.J. Dong, Y.D. Li, *Angew. Chem. Int. Ed*. **42,** 3027 (2003).
45. J. Du, L.Q. Xu, G.F. Zou, L.L. Chai, Y.T. Qian, *Mater. Chem. Phys*. **103,** 441 (2007).
46. Y. Jiang, X.M. Meng, W.C. Yiu, J. Liu, J.X. Ding, C.S. Lee, S.T. Lee, *J. Phys. Chem. B* **108,** 2784 (2004).

47. L. Jin, W.C.H. Choy, Y.P. Leung, T.I. Yuk, H.C. Ong, J.B. Wang, *J. Appl. Phys.* **102,** 044302 (2007).
48. Y.P. Leung, W.C.H. Choy, I. Markov, G.K.H. Pang, H.C. Ong, T.I. Yuk, *Appl. Phys. Lett.* **88,** 183110 (2006).
49. S.V. Pol, V.G. Pol, J.M. Calderon-Moreno, S. Cheylan, A. Gedanken, *Langmuir* **24,** 10462 (2008).
50. M.S. Selim, R. Seoudi, A.A. Shabaka, *Materials Letters* **59,** 2650 (2005).
51. H.P. Klug, L.E. Alexander, J. Wiley & SonsX-Ray Spectrometry, **4,** A18 (1974).
52. J. Du, L. Fu, Z. Liu, B. Han, Z. Li, Y. Liu, Z. Sun, D. Zhu, *J. Phys. Chem. B* **109,** 12772 (2005).
53. H.J. Watzke, J.H. Fendler, *J. Phys. Chem.* **91,** 854 (1987).
54. N.S. Belova, A.A. Rempel, *Inorganic Materials,* **40,** 3 (2004).
55. B. Andersson, J. Gjonnes, A.R. Forouhi, *J. less-Common Met.* **61,** 273 (1978).
56. W. Wei-Yu, J.N. Schulman, T.Y. Hsu, F. Uzi, *Appl. Phys. Lett.* **51,** 710 (1987).
57. B. Nelson, D.P. Riley, *Proc. Phys. Soc.* **57,** 160 (1945).
58. L.V. Azaroff, M. Buerger, *The Powder Method in X-ray Crystallography* (New York: McGraw-Hill: 1958).
59. M.A. Russak, J. Reichman *J. Electrochem Soc.* **129,** 542 (1982).
60. O.V. Krylov, Catalysis by Nonmetals (Academic Press, New York), 258 (1970).
61. D.W. Palmer, *Properties of the II − VI compound semiconductors* www.semiconductors.co.uk (2008).
62. N. Hizem, A. Kalboussi, R. Adhiri, A. Souifi, *J. Microelectron.* **38,** 496 (2007).
63. M.A. Hines, P. Guyot-Sionnest, *J. Phys. Chem, B* **102,** 3655 (1998).
64. F.T. Quinlan, J. Kuther, P. Stroeve, *Langmuir* **16,** 4049 (2000).
65. S. Bhaskar. P.S. Dobal, B.K. Katryar, *J. Appl. Phys.* **85,** 439 (1999).

to Contents, page 1

J. Nano- Electron. Phys.
3 (2011) No1, P.203-211

PACS numbers: 47.70. – n, 68.49.Uv

MOCVD OF COBALT OXIDE USING CO-ACTYLACETONATE AS PRECURSOR: THIN FILM DEPOSITION AND STUDY OF PHYSICAL PROPERTIES

S.M. Jogade, P.S. Joshi, B.N. Jamadar, D.S. Sutrave

O D.B.F. Dayanand College of Arts and Science,
Solapur, 413003, Maharashtra, India
E-mail: sutravedattatray@gmail.com

Metal Organic Chemical Vapor Deposition (MOCVD) is the deposition method of choice for achieving conformal uniform (composition and thickness) continuous thin films over the micron geometry topology necessary for implementing advanced devices. Thin films of cobalt oxide were prepared by MOCVD technique on alumina substrate using a cobalt acetylacetonate as precursor. The thin films of cobalt oxide were deposited on alumina substrate by MOCVD at four different temperatures viz 490 °*C,* 515 °*C,* 535 °*C,* 565 °*C. The as deposited samples are uniform and well adherent to the substrate. Thickness of the cobalt oxide film is maximum at temperature* 535 °*C. The crystalline and phase composition of films were examined by X-ray diffraction. The XRD reveals the crystalline nature with cubic in structure for all the samples. The surface morphology of the films were studied by scanning electron microscopy. The SEM image shows well defined closely packed grains for all the samples. The hexagonal shape of grains are observed for sample at temperature* 515 °*C. Raman spectroscopy shows Fm3m,* 225 *space groups for cobalt oxide thin films deposited on alumina substrate.*

Keywords: *MOCVD, THIN FILM, XRD, SEM IMAGES, RAMAN SPECTRA.*

(Received 04 February 2011, in final form 04 February 2011)

1. INTRODUCTION

In response to the changing global landscape, energy has become a primary focus of the major world powers and scientific community. There has been great interest in developing and refining more efficient energy storage devices. One such device, the supercapacitor, has matured significantly over the last decade and emerged with the potential to facilitate major advances in energy storage. Supercapacitors, also known as ultracapacitors or electrochemical capacitors, utilize high surface area electrode materials and thin electrolytic dielectrics to achieve capacitances several orders of magnitude larger than conventional capacitors [1]. Electrodes for such supercapacitors are mainly made using conductive polymers or metal oxides. The metal oxide based supercapacitors have attracted increasingly more attention due to their high specific capacitance, long operation time and high output.

Metal oxides have many interesting properties that result in various important applications [2]. Transition metal oxides (TMO), a sub group of metal oxide are those oxides in which the cation has incompletely filled d or f shells [3, 4]. Tremendous efforts have been devoted in recent years to study these metal oxides as anomalous behavior observed in these materials.

to Contents, page 1

Consequently it has become increasingly important to understand them in terms of their magnetic, electrical and optical properties. Some of the applications of the transition metal oxides (CaO, NiO, CuO) which have generated lots of interest among the research groups all over the world include superconductivity in electronics, electrocromism in smart windows and electrochemical properties in micro batteries and high density batteries [5].

Among the various metal oxides, cobalt oxides have been extensively investigated because of their potential applications in many technological fields, as well as those of cobalt nanoparticles films obtained by a number of techniques [6-9]. High quality magnetic films of cobalt oxide based alloys are currently used in magnetic heads and magnetic RAM. For example simple binary alloys produce high quality films for magnetic recording industry. On the other hand cobalt oxide based ceramics have attractive magnetic properties and their films and multilayer have been studied for decades and still motivate serious research and development efforts [10-13]. In this paper attempts are made to deposit the thin film of cobalt oxide using metal organic chemical vapor deposition technique (MOCVD). For this the precursor cobalt-actylacetonate is used.

2. BLOCK DESCRIPTION OF MOCVD SYSTEM

Block diagram of typical metal organic chemical vapor deposition system is shown in fig. 1 which consists of following sub units:

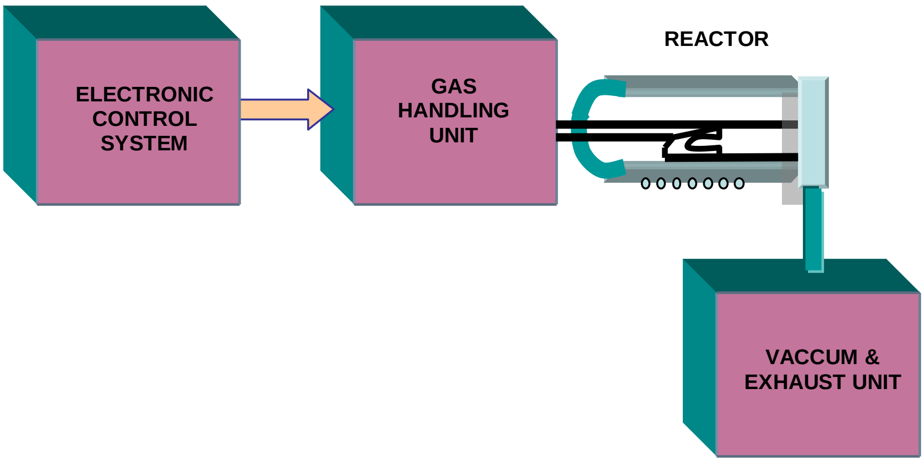


***Fig. 1** – Metal organic chemical vapor deposition system*

2.1 Gas handling unit

The gas handling system performs the functions like, mixing and metering of the gas that will enter into the reactor. Timing and composition of the gas entering the reactor will determine the epilayer structure. Leak-tight of the gas panel is essential, because the oxygen contamination will degrade the growing films' properties. Fast switch of valve system is very important for thin film and abrupt interface structure growth. Accurate control of flow rate, pressure and temperature can ensure the stable and repeat.

to Contents, page 1

2.2 Reactor

A reactor is a chamber where the deposition process is carried on. The chamber is composed of reactor walls, a liner, gas injection units, and temperature control units.

2.3 Vacuum and exhaust unit

Pump and pressure controller is main part of this unit which will control the pressure growth. It is mainly used for low pressure growth. To handle large gas load the pump must be designed in proper manner.

2.4 Electronic control system

This system contains some electronic circuits, which controls various parameters like temperature, rate of flow of oxygen, argon gases, pressure in the reaction chamber etc.

3. THE GROWTH OF THIN FILM BY MOCVD

The vapor pressure is an important consideration in MOCVD, since it determines the concentration of source material in the reactor and the deposition rate. First the metal organic sources and hydrides inject to the reactor. The sources are mixed well inside the reactor and transfer to the deposition area. At the deposition area, high temperature result in the decomposition of sources and other gas-phase reaction, forming the film precursors which are useful for film growth and by-products. Then film precursor's transport to the growth surface, the film precursors absorb on the growth surface, the film precursors diffuse to the growth site. At the surface, film atoms incorporate into the growing film through surface reaction. The by-products of the surface reactions absorb from surface. The by-products transport to the main gas flow region away from the deposition area towards the reactor exit.

4. STEP BY STEP PROCESS IN DEPOSITION

4.1 Cleaning of substrates

The substrate used for depositions is alumina Al_2O_3. Prior to each deposition, the substrates, the substrate holders and the reaction chamber were scrubbed by using detergent, distilled water, trichloroethylene, acetone, ethyl alcohol and distilled water, respectively. The substrates were washed with 1:1 Hydrochloric acid and double distilled water. Then it is cleaned by ultrasonic cleaner for ten minutes. Again washed with double distilled water. 50 ml of acetone was boiled, and then with the vapors of acetone substrate were cleaned till entire acetone. After this 25 ml of trichloroethelene was boiled and substrate was cleaned in these vapors, and kept in air-tight container. This process was used to dislodge the dirt on the glass, and ensure that hydrocarbon and grease were removed from the substrate and also to ensure that the substrate surfaces were free from surface contamination and defects [4].

4.2 Preparation of precursor Co-acetylacetonate

The 2.4-pentanedione (acetylacetonate) 40 ml was added slowly to a solution of 16.0 gm of sodium hydroxide in 150 ml of water and kept at a temperature

to Contents, page 1

bellow 40 °C The yellow solution was added drop wise to a solution of 47.6 gm of cobalt (II) chloride hydrate ($CoCl_2.6H_2O$) in 250 ml of water and stirred vigorously. The resulting orange precipitate was filtered in a large Buchner funnel and washed with about 500 ml of water until the washing was colorless The moist solid was then dissolved in hot mixture of 400 ml ethanol and 250 ml of chloroform. The red solution was allowed to cool slowly to room temperature and then further cooled in ice. The orange needles were suction filtered and washed with cold 95 % ethanol and air dried [4]. The reactions are given bellow:

$$HC_5H_7O_2 + NaOH \longrightarrow NaC_5H_7O_2 + H_2O$$

$$2NaC_5H_7O_2 + CoCl_2 + 6H_2O \longrightarrow Co(C_5H_7O_2)_2 + 2NaCl + 6H_2O$$

4.3 Deposition conditions

The thin films of cobalt oxide were deposited on alumina substrate by MOCVD technique. The thin films of cobalt oxides are deposited at four different temperatures viz. 490 °C, 515 °C, 535 °C, 565 °C. The different deposition conditions and parameters are as shown in table 1.

***Table 1** – Deposition parameters and conditions*

Sr.No.	Parameters	Conditions
1	Precursor used for deposition	Cobalt acetylacetonate
2	Substrate used for deposition	Al_2O_3
3	Purging gas	Argon
4	Purging time	30 min
5	Reacting gas	Oxygen
6	Time of reaction	60 min
7	Base pressure	0.06 T
8	Purging gas pressure	0.21 T
9	Deposition pressure	10.00 T
10	Temperature of Vaporizer	185 °C
11	Line temperature	200 °C
12	Temperature of substrate	515 °C
13	Carrier gas (Argon) flow rate	9 %
14	Reacting gas (O_2) flow rate	5 %

5. PHYSICAL PROPERTIES OF THE SAMPLE

The thin films of cobalt oxides are deposited at four different temperatures viz. 490 °C, 515 °C, 535 °C, 565 °C. All the samples deposited on the alumina substrate are well adherent and grayish in color. The thicknesses of the samples were calculated by weight difference method. The table 2 shows the variation of thickness with temperature. Thickness of the cobalt oxide film is maximum at temperature 535 °C.

to Contents, page 1

Table 2 *– Variations of thickness of cobalt oxide thin film at various temperatures*

Sr. No.	Temperature, °C	Thickness, mkm
1	490	0.04205
2	515	0.05600
3	535	0.11120
4	565	0.01905

6. STRUCTURAL ANALYSIS BY XRD

The X-ray diffraction patterns were obtained for all these samples by using Bruker D8 advanced instrument with source CuKα_1 with $\lambda = 1.5406$. The angle-2θ is varied in the range between 10° to 90°. The fig. 2 shows number of peaks observed for various temperatures. All the samples are crystalline in nature with cubic in structure. The as deposited samples show dominating peaks these data are compared with the JCPDs-ICDD data no. 78 - 1970 and JCPDs-ICDD data no.78-0431 [14]. The cobalt oxide films obtained on alumina substrate shows more number of highly intensed peaks. The peak corresponding to a plane (311) of Co_3O_4 is most prominent which is observed for all the four samples and same is confirmed with JCPDS-ICDD data no. 78-1970 for Co_3O_4 [14] Few peaks corresponding to the CoO are also observed. The intensity patterns are more or less similar in all the cases. The peak corresponding to plane (222) shows less intensity as compared to other peaks and same is confirmed with JCPDS-ICDD data no. 78-0431 for CoO [14]. The table 3, shows comparison between intensity of no of peaks observed at different temperatures for angle-2θ with ASTM values. Similar results are obtained by D.Barreca et al. [15].

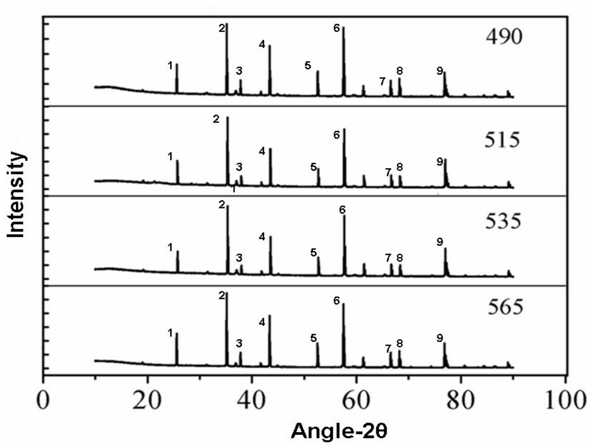


Fig. 2 *– XRD of cobalt oxide thin film deposited on alumina substrate*

to Contents, page 1

Table 3 *– XRD Analysis of cobalt oxide thin films deposited at various temperatures*

Peak No.	Observed data		ASTM data		Plane
	Angle-2θ	Intensity	Angle-2θ	Intensity	
Temperature = 490 °C					
1	26.00	47.2	--	--	--
2	36.25	96.5	36.84	999	(311)
3	38.50	33.2	38.54	080	(222)
4	44.22	68.4	44.80	173	(400)
5	53.35	45.0	--	--	--
6	58.00	88.6	59.34	210	(511)
7	66.50	33.3	65.22	305	(440)
8	68.60	34.0	68.61	002	(531)
9	77.54	45.0	77.52	108	*(222)
Temperature = 515 °C					
1	25.75	45.5	--	--	--
2	36.60.	94.5	36.84	999	(311)
3	38.00	31.0	38.54	080	(222)
4	43.25	61.5	42.38	999	*(200)
5	53.00	43.3	--	--	--
6	58.10	77.2	59.34	210	(511)
7	67.30	37.5	68.61	002	(531)
8	68.50	31.5	68.61	002	(531)
9	77.50	52.0	77.52	108	*(222)
Temperature = 535 °C					
1	26.40	45.5	--	--	--
2	36.65	97.5	36.84	999	(311)
3	38.50	31.0	38.54	080	(222)
4	44.00	62.0	44.80	173	(400)
5	52.50	39.2	--	--	--
6	58.50	77.7	59.34	210	(511)
7	67.25	33.2	68.61	002	(531)
8	69.00	32.0	69.73	001	(442)
9	78.00	46.0	78.39	032	(622)
Temperature = 565 °C					
1	26.25	48.0	--	--	--
2	36.28	97.5	36.84	999	(311)
3	37.75	33.0	38.54	080	(222)
4	44.00	74.8	44.80	173	(400)
5	53.00	43.5	--	--	--
6	58.00	88.0	59.34	210	(511)
7	67.00	33.5	68.61	002	(531)
8	68.50	35.0	68.61	002	(531)
9	77.50	44.8	77.52	108	*(222)

to Contents, page 1

7. MORPHOLOGICAL CHARACTERISTICS

The SEM images were obtained from ESEM Quonta 200 instrument. The SEM images obtained for cobalt oxide thin films on alumina substrate shows more packed grains with increase in size of grain. The hexagonal shape of grains are observed for sample at temperature 515 °C. Above this temperature the grain size found to be decreased. At temperature 565 °C the different shapes of grains are observed. The patterns are shown in fig. 3. The table 4 shows average grain size of deposited thin films observed from observed from SEM images. This is confirmed by Mordi et al. and Nygirnyi et al. [4, 16].

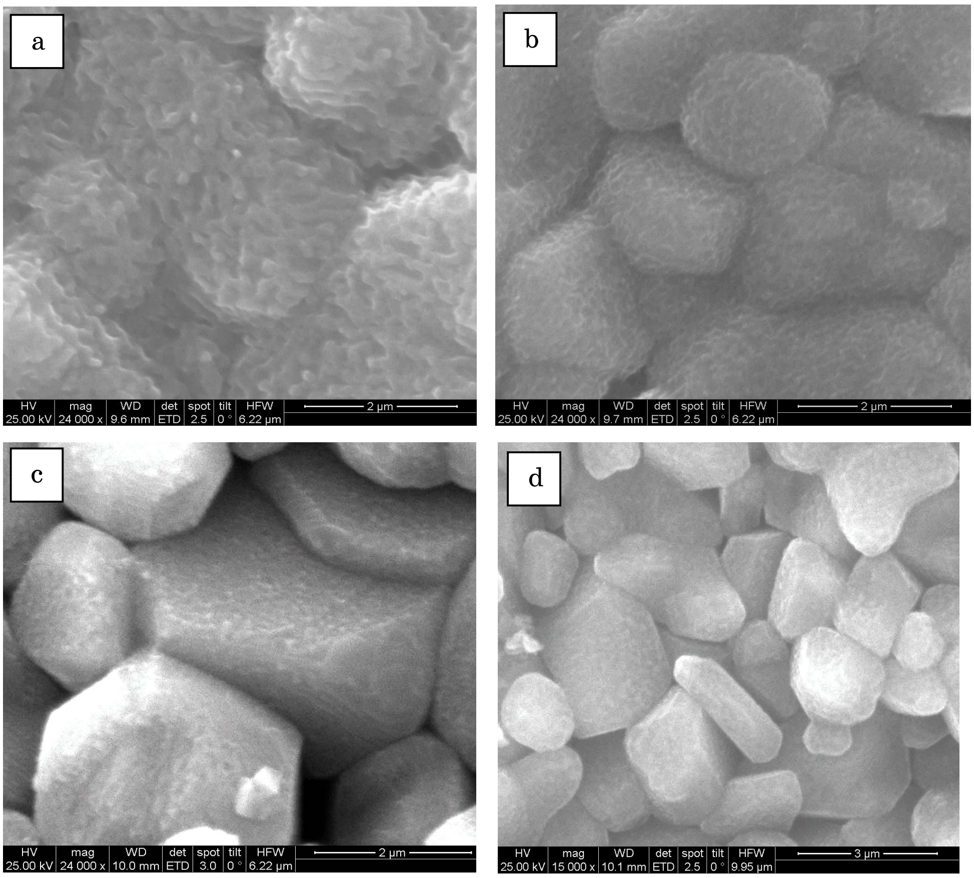


***Fig. 3** – SEM images of cobalt oxide thin film deposited on alumina substrate at T = 490 °C (a), T = 515 °C (b), T = 535 °C (c) and T = 565 °C (d)*

8. RAMAN SPECTRA ANALYSIS

The RAMAN Spectra for the cobalt oxide thin films are obtained with the help of "NSOM" instrument. The similar characteristic is observed for all the samples deposited on alumina substrate. Raman spectroscopy shows Fm3m, 225 space groups for cobalt oxide thin films deposited on alumina substrate. The Fig. 4 shows RAMAN Spectra for the cobalt oxide thin films deposited on alumina substrate.

to Contents, page 1

Table 4 *– Average grain size observed from SEM images for various temperatures*

Sr.No.	Temperature, °C	Average grain size, mkm
1	490	2.82
2	515	2.35
3	535	3.26
4	565	2.55

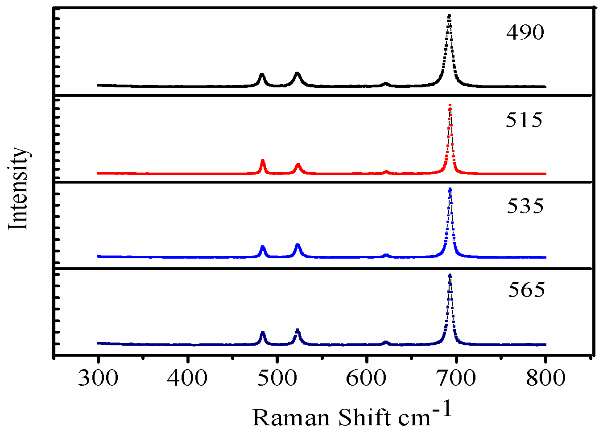


Fig. 4 *– Raman spectra of cobalt oxide thin film deposited on glass substrate*

9. CONCLUSION

The MOCVD technique is most suitable to deposit good quality thin films of cobalt oxide from a cobalt acetylacetonate precursor. The as-deposited samples are well adherent to the substrates. The samples are a crystalline in nature with cubic in structure. The SEM images shows well developed packed grains of cobalt oxide. Raman spectroscopy shows Fm3m, 225 space groups.

REFERENCES

1. M.B. Bever, Encyclopadia of material Sciscиencea and Engeneering (Pergamon: Oxford: 1986).
2. J.H. Richter, *Electronics properties of metal oxide films studied by core level spectroscopy*, Digital Comprehensive Summaries of Uppsala Faculty of Science and Technology, 228, 69.
3. M.S. Halper, J.C. Ellenbogen, *Supercapacitors: A Brief Overview*, (Virginia: MITRE: 2006).
4. S.K. Ray, G. Sasikala, *J. Mater. Sci. Technol.* **25**, 85 (2009).
5. C.M. Lampert, C.G. Granqvist, *Large-Area Chromogenics: Materials and Devices for Transmittance Control*, *Vol. IS4* (SPIE Optical Engineering Press: Bellingham: 1990).
6. L.M. Apatiga, V.M. Castano, *Thin Solid Films* **496**, 576 (2006).
7. L. Horng, G. Chern, P.C. Kang, D.S. Lee, *J. Magn. Magn. Mater.* **270**, 389 (2004).
8. S.M. Zhou, S.J. Yuan, L. Sun, *J. Magn. Magn. Mater.* **286**, 211 (2005).
9. L.P. Barbosa, H. Takishi, R.N. Faria, *J. Magn. Mater.* **268**, 132 (2004).

to Contents, page 1

10. A.U. Mane, K. Shallini, A. Wohlfart, A. Devi, S.A. Shivshankar, *J. Cryst. Growth* **240,** 157 (2002).
11. K. Shalini, A. U.Mane, S.A. Shivshankar, M. Rajeshwari, *J. Cryst. Growth* **231,** 242 (2001).
12. M.A. Paranjape, A.U. Mane, A.K. Roychoudhary, K. Shalini, S.A. Shivshankar, B. Chakravarty, *Thin Solid Films* **413,** 8 (2002).
13. A.U. Mane, S.A. Shivshanker, *J. Cryst. Growth* **254,** 368 (2003).
14. JCPDS data file No. 78-(1970), 78-0431.
15. B. Barreca, C. Massignan, S. Daolio, M. Fabrizio, C. Piccirillo, L. Armelao, E. Ondello, *Chem. Mater.* **13,** 588 (2001).
16. V.M. Nagirny, R.D. Apostolova, A.J. Baskevich, P.M. Litvin, E.M. Shembel, *J. Rus. Appl. Chem.* **75**, 905 (2002).

to Contents, page 1

Privacy on a Network

Juan Manuel Fernández López

This article offers an overview of the European and Spanish law on the protection of the personal privacy with respect to the use of the information technologies. It describes the characteristics of the Internet with respect of these Rights. It also summarises the recommendations and documents elaborated by the Data Protection Agency of Spain and other European organisms, as well as the Spanish security regulations.

***Keywords**:* protection of personal privacy, protection of personal data, data collection, Internet, European law, Spanish law, Group 29.

Privacy is a fundamental right of human beings. The Universal Declaration of Human Rights proclaimed by the General Assembly of the United Nations on 10 December 10 1948 establishes in article 12 that "*(…) No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks*." For its part, the European Convention for the Protection of Human Rights and Fundamental Freedoms of the Council of Europe in 1953 states in Article 8 that "*Everyone has the right to respect for his private and family life, his home and his correspondence*." In both texts one must consider both the right to privacy and its embodiment in the protection of personal computerized data.

The extension of the right to privacy with regards to computerized processing of personal data is defined in the Convention for the Protection of Individuals with Regard to the Automatic Processing of Personal Data and also Treaty 108 of the Council of Europe, dated January 28, 1981. In its preamble this treaty indicates that in order to "*(…) achieve greater unity between its members, based in particular on respect for the rule of law, as well as human rights and fundamental freedoms (…) it is desirable to extend the safeguards of everyone's rights and fundamental freedoms, and in particular the right to respect for privacy, taking account of the increasing flow across frontiers of personal data undergoing automatic processing*."

As it pertains to Spain the Spanish Constitution in Article 18 makes a positive formulation of this right to family and personal privacy and in Clause 4 sanctions legislature to "*(…) limit the use of information, to guarantee personal and family honour, the privacy of citizens, and the full exercise of their rights*." This constitutional mandate was complied with the promulgation of the Basic Law 5/1992, 29 October, the regulation of the automated processing of personal data[1] that was in effect until January of this year, from which point it was superseded by a new law: the Basic Law 15/1999, of 13 December 1999, of protection of personal data[2]. The reason for the approval of the new Law has been, fundamentally, to complete the transposition to Spanish law of all the principles contained in Directive 95/46/CE, of 24 October 1995, relating to the protection of physical individuals as it pertains to the processing of personal data and to the free circulation of this data. Both legal texts sanctify a fundamental series of principles that must always be observed in the processing of data. Of these, three deserve special attention according to my criteria. Thus, before the collection of data the citizen must be informed of the compulsory reason, or lack thereof, for providing the information, of the liabilities database of the purpose for which the data is collected. Furthermore, for the processing of data as a general practice, the affected party must provide his consent and, with exceptions as outlined, the data may not be used for purposes other than the original intention of collection and may only be used with the citizen's consent.

Additionally, and with the purpose of adapting principles of data protection and privacy to a sector that presents particular specifications, Directive 97/66/CE was published on 17 December relating to the processing of personal data and to the protection of privacy in the telecommunications sector. This was incorporated into the Spanish Right by means of Law 11/1998, General of Telecommunications and the Royal Decree 1736/1998.

Juan Manuel Fernández López belongs to the Spanish Judicial branch as a judge. Since 1998 he is director of the Data Protection Agency. He is vice-president of the Court of Defence of Competence, and professor of Commercial Law at the University Complutense of Madrid. He belongs to the International Association for the Protection of the Industrial Property, and to the Advisory Committee of the Center for Studies in European pharmaceutical law. Vice-president of the Institute of Industrial Law and Ethics, he is also member of the Royal Academy of Jurisprudence and Legislation, member of the General Commission of Codification and Director of "Comunicaciones IDEI", law magazine.
<https://www.ag-protecciondatos.org>

1. LORTAD: Ley Orgánica de Regulación del Tratamiento Automatizado de los Datos de Carácter Personal.
2. LOPD: Ley Orgánica de Protección de Datos de Carácter Personal

Once examined succinctly, we should ask ourselves, in the normative framework that guarantees the right to personal and family privacy, if the special characteristics of Internet, that is to say its global and non-territorial qualities, make it a special field where these rules do not apply. The answer is that a legal vacuum does not exist. Directive 95/46/CE, as well as Directive 97/66/CE and the national Laws, are applicable to the Internet. This further implies that the principles of data protection, as well as the rights of the pertaining citizens, also apply when the processing occurs on the Internet.

One of the fundamental principles of the protection of data is that the data obtained in any situation shall be limited to what is deemed necessary and pertinent for the purpose. A characteristic of a telecommunications network, and of the Internet in particular, is its capacity to generate a huge quantity of transactional data.

Whenever the Internet is accessed, a digital trace is left, so that, with a growing number of our day-to-day tasks being carried out online, information is ever increasing with regard to our occupations, tastes and preferences.

Taking all of this into consideration, the threat to our privacy is not only derived from the existence of large quantities of personal data on the Internet, but also from the development of logical media capable of searching the network and compiling all of the available data on a specific person.

Upon accessing and extracting data from said media, detailed biographical details of a random person can be compiled covering all of the groups in which the said person has participated: the address and phone number of the selected person, as well as discover where they were born, where they studied, their profession, and their current place of employment. It could also detail their interest in amateur theatre, their favourite type of beer, their preference in the matter of restaurants and vacation resorts, and their opinions about various topics. In the United States, we already see the existence of many Internet companies that commercialise these "search services."

Due to all this the protection of the privacy of Internet users has been a constant source of concern to the Data Protection Agency such that it had as early as 1997, elaborated and published some Recommendations for Internet Users, whose more important points can be summarised as follows:

- When providing personal data to any organisation (access providers, content providers, e-commerce vendors, etc.) be conscious of whom you are providing it to and for what purpose.
- Try to ascertain the policy of your providers as well as list and directory administrators as they pertain to a sale, exchange or rent of the data with which you supply them.
- Request that your personal data not be attached to your Internet login.
- Distrust any request for information if the data being requested is excessive for the purpose it is being collected or unnecessary for the service being rendered.
- Keep in mind that upon entering your e-mail address in a directory, distribution list or newsgroup, said address can be collected by third parties to be utilised with a different purpose, for example, to send you unsolicited marketing e-mail.
- While navigating the Internet, be conscious that the Web servers that you visit may record the pages you access as well as the frequency and the topics or subjects that you search, although you may not be informed of it. Similarly, your belonging to specific newsgroups and distribution lists may contribute to the development of more or less detailed profiles about you.
- In case you do not wish to leave a trace of your Internet activities, you can use programs that preserve anonymity, such as the use of electronic money or servers that provide this service.
- Whenever possible, use the latest versions of Internet browsing programs, since they incorporate more and more security measures. Consider the possibility of activating in said programs the options that alert you upon the undesired exchange of data and do provide any data which you do not wish to make public (for example, e-mail address, name, last names, etc.)
- Do not conduct commercial electronic transactions through servers which have been deemed "insecure" or unreliable. Consult the browser manual to ascertain how a connection with a secure server has been established.
- Remember that there are electronic money systems that preserve the anonymity of your Internet purchases.
- Use the security mechanisms available to you to protect your data from undesirable access. The most reliable means of obtaining security is through encryption.
- Unless reliable mechanisms are being used, such as authentication and certification (digital signatures, public key encryption, etc.) do not blindly trust that the person or organisation to which you send a message is the one it claims to be and that the content of the message has not been altered, although this is true in the vast majority of cases.
- Every time personal data is requested that you are not legally bound to supply, weigh the benefits you are going to receive from the organisation collecting it versus the possible risks of irregular use of the data.
- If in any doubt about the legality of the use of personal data, contact the Data Protection Agency.

These recommendations, which were pioneering in their time and were recognised as such by the European Commission, have been complemented in recent years by a series of documents issued by the Group for the Protection of Individuals with respect to the processing of personal data (Working Group Article 29).

The Working Group Article 29, with the declaration of "The processing of data on the Internet" intends to establish clearly that Directive 95/46/EC, as it pertains to the protection of personal data, and Directive 97/66/EC, as it pertains to the Telecommunications sector, apply without a doubt to the Internet. Nevertheless, with the objective of studying these topics, an *ad hoc* subgroup has been created: the Operative Internet Group.

Thanks to the work of this subgroup, some important documents have already been approved:

- Recommendation 2/97: Supports the Budapest-Berlin Manifest, of the International Working Group, regarding the protection of data in the Telecommunications sector.

- Recommendation 3/97: Anonymity on the Internet.
- Opinion 1/98: Report regarding project P3P of the Consortium of the W3C.
- Working document: “The processing of data on the Internet.”
- Recommendation 1/99: Invisible processing on the Internet carried out by means of software and hardware.
- Recommendation 2/99: The respect of privacy in the interception of telecommunications.
- Recommendation 3/99: Storage of traffic data by Internet Service Providers to fulfil legislation.

These documents especially stress that maintaining the possibility of anonymity is fundamental so that the privacy of the individuals may be subjected to the same protection online as offline, although it is also recognised that anonymity is not always possible. The circumstances in which anonymity is possible, fundamental rights or privacy and freedom of expression must be weighted against the important objective of maintaining public order including the prevention of crime. The legal restriction that may be imposed by the Government with regards to the right of maintaining anonymity or of the technical means by which to produce the same (for example, availability of encryption products), must always be provided and be limited strictly to what is necessary to protect a specific general interest in a democratic society. The balance reached in relation to earlier technologies must be preserved much as possible when it pertains to the services offered through the Internet.

Additionally, the software and hardware industries are encouraged to work on products that protect privacy and that contain the necessary tools to comply with the European rules on data protection. Personal data can be legitimately processed when the affected party is informed and is consequently aware of the processing. Because of this, the Working Group is particularly concerned about all those processing operations that are currently carried out by software and hardware on the Internet without the knowledge of the affected party and as such are not perceived by the affected party.

On the other hand, one must explicitly inform the Internet user what data is being collected, whether explicitly or implicitly, thus giving him the opportunity to object to its processing.

Provision has been made for the Operative Group to continue work for the remainder of the year 2000 and, as a result, it will be able to provide a general document that will tackle all the aspects which are relevant to the processing of personal data on the Internet. The document, which will be submitted to the plenary Group of Article 29, is expected to serve as a reference on the opinion of controlling authorities and to be distributed to different members of the European Union. Additionally, it will constitute a very useful tool in the national arena as a guideline to harmonise the approaches to the different problems that may arise.

Similarly, in 1999, the Council of Europe approved Recommendation (99) 5 regarding the protection of privacy on the Internet, in which this problem is approached from the double perspective of the user and of Internet service providers.

A very important aspect of guaranteeing the privacy of the citizen on the Internet is security. Recently, we have all been informed by the mass media of a series of situations that have threatened the confidentiality, the integrity and the availability of a good number of computer systems connected to the Internet. These incidents have re-ignited the interest in one of the elements which must be further developed if we want citizens to trust the Internet as a secure means of joining the information society without the risk of compromising their privacy, their finances or both.

It is particularly opportune to mention the Basic Law 15/1999, with regards to the protection of personal data: the adoption of measurement and organisational techniques that guarantee the security of personal data and avoid their alteration, loss, or unauthorised processing or access. I am referring to Royal Decree 994/1999, which approves the Regulation of security measures on automated databases that contain personal data. The Regulation determines the means by which different techniques that guarantee the confidentiality and integrity of the information with the purpose of preserving the honour, the family and personal privacy and the full exercise of the personal rights regarding their alteration, loss, unauthorised processing or access.

The security measures that are established are the basic security requirements that all databases containing personal data must fulfil, they do not prejudice the establishment of special measures for those databases that require a greater degree of protection due to the nature of the information they hold. For these, the Regulation has established three levels of security (basic, medium and high) corresponding to the nature of the information and the relative consequences for privacy of citizens should a violation of the necessary confidentiality and integrity of the database information occur.

The fundamental objectives of the Regulation are, therefore:

- To create a general framework to help to develop and carry out organisational and measurement techniques that increase the guarantees of the processing of personal data.
- to make all the relevant parties aware of the requirements for processing personal data and the possible risks related to its use as well as the need to put security measures to place.
- To complement the security measures with management-related administrative and organisational measures.
- To facilitate the classification of databases and processing of data in relation to the risks that they present and the nature of the data to be protected.
- To ensure an appropriate level of security. That implies a balance between the risks, existing technical solutions and the cost of implementation of solutions.
- To ensure a periodic evaluation of the security measures put into effect.

As I have already indicated, all of the legislation is in effect with regards to the computer systems connected to the Internet. As such, personal data should receive adequate protection by means of the effective implementation of the provisions of the Regulation of Security.

(English by Mike Andersson and Adam David Moss)

to Contents, page 1

Network Privacy: State of the Art

Félix Herrera Priano

Security, like many other network aspects, has a physical component (security of the hardware) and a logical component (organisation and enterprise security policies). Only a correct combination of both allows an approximation to an effective network security model. We also cover other aspects such as centralised and distributed security models at all security levels by means of firewall systems.

Keywords: security for hosts, firewall systems, security policies, private virtual networks, intruder detection systems, hackers.

1 Introduction

The widespread adoption of computer networks and the World-Wide Web to interconnect information resources of all types has not only considerably increased the scope and variety of the risks to information stored on such systems, it has also exposed the weaknesses in many of the security measures currently taken to limit the impact of such risks.

The social and technological revolution entailed the development of Internet technologies, which was initially born out of a number of networks for the exchange of information among researchers collaborating on joint projects, or sharing results. While in this early work, there was not much concern about data privacy or security matters, even so, it was thought not to be recommended to transfer sensitive or classified documents over the Internet. Given the origins of the Internet in the U.S. Department of Defence's need, in 1968, to interconnect universities and research centres that collaborated in some way or other with the U.S. Armed Forces, such security considerations were clearly apparent.

With links to the business world, all kinds of information now travels on the network, including complex transactions requiring some security measures to guarantee a given level of confidentiality.

This raises a number of questions: How can inappropriate access to applications and stored information be prevented? How can we reliably assess that the sender and the addressee of a communication really are who they claim to be? Or, how can we guarantee an audit trail to prevent a sender denying having sent a message or an addressee denying its receipt?

We have seen in increase in both the variety and volume of users on the Internet with various applications such as education, research, business intelligence and business-to-business linking. In the midst of all this, there has also been an increase in abuse, violating the privacy and ownership of resources and systems. Hackers, lammers, crackers, viruses have become enemies faced by network and system administrators.

The very complexity of current networks hampers the detection and resolution of the many security problems that occur. Security is such a fundamental requirement for the design of networks that it must be adequately specified. We must decide between distribution and concentration and take advantage of the technologies and advanced devices that allow greater degree of effectiveness in matters of information security.

2 Host and Network Security

The security or, more pertinently, the lack of security on the Internet, is undoubtedly a major problem and has given rise to financial loss in recent years in government, public administration and business. Even to establish the fundamental parameters that govern world network security is complicated. We can break down information security in two main categories: Host security, and network security.

The best way to implement a security project is to develop strategies in combination in both these areas, even though it may often be more complicated, especially in large organizations.

Host security includes the combination of techniques and tools that provide protection for an individual computer: correct operating system configuration, backup procedures, file encryption, antivirus programs, audit programs etc. As well as the cryptographic techniques and tools it is important to stress that a key component in system protection lies in the systematic and continuous application of precautionary measures by network administrators.

Network security, the second category, comes into play in interconnected systems, involving aspects such as network authentication, system firewalls, Intrusion Detection Systems (IDS), monitor programs etc.

Additionally, within network system security we can identify two main trends:

Félix Herrera Priano is Professor at the University of La Laguna and researcher of the Communications and Teledetection Laboratory of the ULL. He graduated as telecommunications engineer at the University of Las Palmas. Postgraduate Research at Queen's University, UK. He has worked in different fields related with the implementation of information and communications technologies, corporate security, smart cards and new technologies. Author of various papers and books. <fpriano@ull.es>

- *Protection of transfer systems:* In this case, the service administrator assumes responsibility for guaranteeing the transfer is secure in a way transparent to the user. Examples of this type of network would be the establishment of a secure level of transfer, or the installation of firewalls, that defend access to a protected area of a network.
- *Secure point-to-point applications:* This is the case of a digitally signed document, or, for example, an e-mail the content of which has previously been protected by means of some procedure. Although the final responsibility for the security of a message lies with the user, it is reasonable to expect that a tool provided for such security in his or her organization should be employed.

3 Security Plans and Policies

The term security policy encompasses among other things the procedures to be followed and the relationships between users and systems to ensure these are used correctly. Security policy is company specific and must be adapted to each company's particular needs. An even more all-encompassing concept would be that of the plan where all necessary points are taken into consideration in a particular organization's implementation. In normal circumstances, security policy should be an embedded component of the organisation's normal standard and procedures.

Developing security policy can become complicated if it also helps determine what we need to protect, how much investment is needed to protect the network's security and who will be responsible for maintaining the protection. Prior to the design of the security policy, The first step is the development of an effective security plan based on an understanding of the threat that connectivity represents for the system. We must identify and quantify these dangers as a function of the number of people that may be affected, as well as of the information that is jeopardized. It may seem obvious that those systems with classified or sensitive information should not be connected directly to the Internet. For example, personal information, medical histories, critical references etc. need to be protected, denying unauthorized access.

Denying access to services can cause many problems, especially if many users are affected or if the main function of an organization depends exclusively on those services.

Let us not forget that it is people and not computers who are responsible for the implementation of security procedures. And it is also people who violate a system's security. As such, a network's security is useless if responsibilities are not clearly delineated.

A security policy for networks should define the roles at each level of the network:

- The responsibility of the network user: One policy might require that the users change their passwords frequently, or that they fulfil a series of requirements. They may also be asked to check their accounts to see if anyone has accessed them, so that possible intrusions may be detected.
- The responsibility of the system administrator: The policies might require that specific security measures be identified, intrusion messages, procedure that are executed every so often to control the actions of each user on each server, etc.
- The correct use of network resources: Rules governing who should use network resources, what they can and cannot do should be implemented. For example, if it is considered that e-mail, file folders and computer logs are subject to control, the network policy should be immediately communicated to the users.
- Actions to be taken upon detection of a security problem: Questions such as: What should be done? Who should be notified? It is very easy to take these things for granted in complex situations. One must specify the exact steps that a system administrator must take when a breach in security is detected.

4 Effective Security: System Firewalls

A great number of Internet security problems could be resolved, or at least reduced, by means of the use of existing well-known techniques and controls such as system firewalls. A firewall significantly improves an organization's level of security, while allowing vital access to Internet services. The concept of firewalls, which emerged as commercial products in 1995, has created a revolution in the information technology industry as a means of centralising network security (Figure 1).

Aside from the more secure and logical method of not connecting, a firewall is undoubtedly the "most effective model that currently exists for the implementation of network security". However, it is wise to remember that most attacks originate from within an organization, where this type of system does not work.

A firewall is a focal point for attempts to ensure security; it helps build a security policy that defines which services and information can be accessed and which implements that policy from the point of view of the network configuration and of other protective measures such as authentication. The main purpose is to control the access into or out of a protected network. As such, an access policy is established in the network forcing

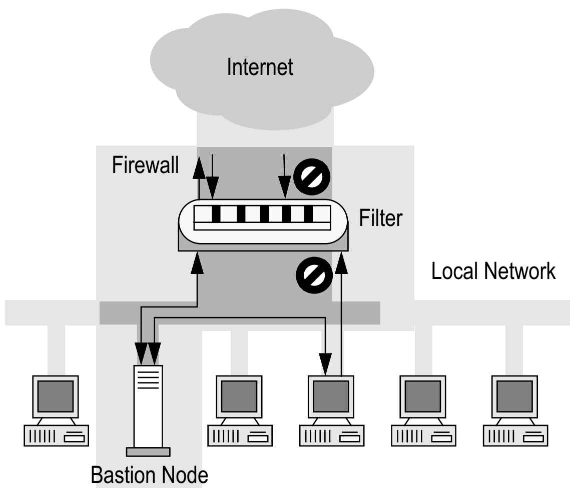


Figure 1: Use of firewalls

to Contents, page 1

all connections to be made through the system firewall, where each attempt at access can be examined and evaluated.

While a system firewall is normally placed at the outermost level of accessibility, such as the connection to the Internet, they can also be located inside an organisation's system to provide protection to smaller groups of hosts or sub networks.

Firewalls are not always required: The majority of workstations and computers may contain information or applications that do not need the level of protection provided by firewalls. Quite often, only a limited number of an organisation's systems contain information and processes, critical to the organisation's functioning.

A way to limit the impact of a firewall on the operations of a network is to use an internal one that isolates the critical systems, allowing the remainder of the network to operate normally without filtering conditions (Figure 2).

Privacy becomes of great importance in maintaining secure organizations, since the information that can be accessed may contain useful hints to the attacker. If all the access to and from the Internet goes through this type of system, access can be registered and valuable information may be provided. It is important to collect statistics regarding the use of the network and to record the activities for a number of reasons. The most important is to find out if such devices resist attempted intrusions and attacks, and to determine if the controls are adequate. The statistics regarding network use are also important as information for identifying network requirements and for risk analysis.

A firewall also supplies the means by which to implement and to impose a network access policy. In effect, a firewall provides access control for both users and services so that a network access policy can be put into effect. Without a firewall, such policy would depend entirely of the cooperation of the network users. While an organization may be able to rely on cooperation from its own users, it cannot, nor should not, be dependent on the users of other networks.

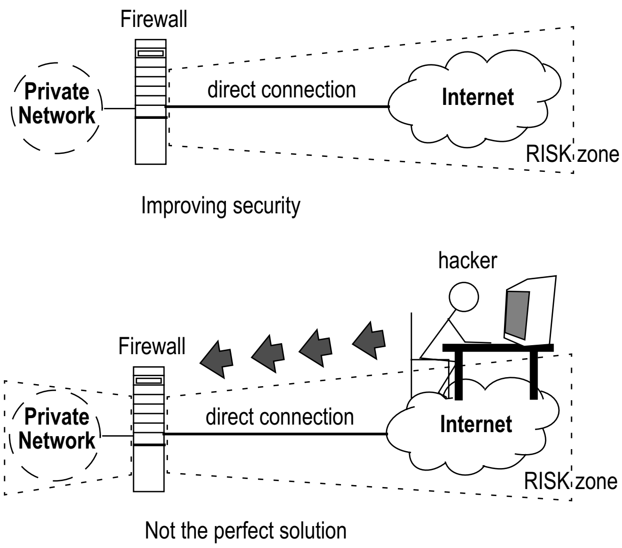


Figure 2: Internal Firewalls

5 Distribution of Control

Another focus of network security is the distribution of responsibilities for small segments of a large network within the organization. Distribution of responsibilities and control can create a web of small networks made up of reliable servers. In this case, the majority of security activities take place on individual systems. The developers of these systems should be familiar with security responsibilities and their contribution to the security of the network should be recognized and appreciated.

The subnetworks are a powerful tool in the distribution of control. When a subnetwork is created, an administrator needs to be appointed to bear the responsibility for the security of that portion of the network with authority to, for example, allocate IP addresses to the devices connected to it. In any case, under this model of security distribution, it is not necessary to abandon a centralized security system. In fact, the establishment of security points in a distributed manner and the maintenance of a single security administrator is currently common practice in many businesses. VPNs, or virtual private networks, are based on this design. In these cases, an interconnection support, such as the Internet, can be used to offer high levels of security to different elements of an organization that may be distributed around the world (VPNs work using communications encryption procedures) (Figure 3).

6 What Does the Future Hold for Us?

In looking at the real state of the art of network security and privacy, we should consider the following: As we already know, security is inversely proportional to the extent of the services offered. Current operating systems include and may require from the user many functions that are normally supplied through the deployment of millions of lines of code. This last, can make us think that systems nowadays can be more vulnerable than that concept of information technology through monolithic systems that offer the user little configuration flexibility.

The best security information is found on the Internet. More and more, security will become a key element of a network. In software applications, errors do not specify themselves, they simply appear. This is the key. As much as we try to develop stable applications, there are too many factors that come into play: we might have to deal with badly designed hardware, an error prone or wrongly configured operating system or even the programming language in which we are developing our code could possibly include security problems. Aside from that, the communication protocols might suffer from the same types of errors. Time has shown us that security problems are and will be present, they are inherent in the very design and, as such, we should continue to adapt ourselves and configure our systems to attempt to develop an effective security model.

Attack techniques change all the time, operating systems are continuously being developed and applications are more and more flexible. In fact, it is becoming increasingly common not only to alter the attack techniques but also to use them in conjunction with each other in order to make detection by IDS, or Intrusion Detection Systems more difficult. It is also true that we have become accustomed to paying the price for the lack of good quality program code by increasing machine resources by adding memory, disk or a more powerful CPU.

There are opposing views on whether it is better to base security systems on widely used operating systems like Windows or Unix or to use more specialised propriety systems. For Unix and Windows, we have some advantages and disadvantages: the fact that we have security platforms based on systems with these characteristics can make them tempting to the attacker but, at the same time, the fact that many groups are involved in discovering and solving the security problems that arise means that 'quick fixes' are usually available. On the other hand, a specialised proprietary system may be less attractive for the hacker because a limited distribution limits the scope of a successful attack. In any case, this type of system can offer the same or greater security problems than the more popular ones.

Another critical security area is the digital signature and the potential revolution this could create. In effect, new systems based on signatures will allow confidential, private and authenticated transactions to be carried out between separate network systems. But in order to gain the most benefit from secure electronic transactions, we must first analyse the internal organizational procedures to be able to undertake worthwhile projects based on these types of technologies. It should not be used, as many institutions erroneously believe, as the solution to an organization's functional problems. When an organization uses work flow procedures with incorrect information, the use of a solution based on digital signatures does not contribute at all to the speed with which the information can be processed.

Currently, security is experienced as an additional layer that we include on our systems and services. More and more, security is a procedure included in the communication protocols and indeed implemented in specific hardware.

It creates a technological conflict, pitting the speed of communications networks against the need to filter information: are available that provide a large bandwidth for data interconnection. What happens is that, due to security problems, it is sometimes necessary to filter or analyse the information transmitted to and from companies. On the one hand, we have been able to get high rates of data transmission without problems, but on the other hand, control elements of increased sophistication and speed for the analysis of information have to be incorporated. For this reason, recent years have seen the profile of the basic network security device, the system firewall, involving more hardware components rather than software, specifically seeking greater speed. Additionally, although they already exist on the market at high prices, the future of interconnection is likely to be through communication devices that directly implement

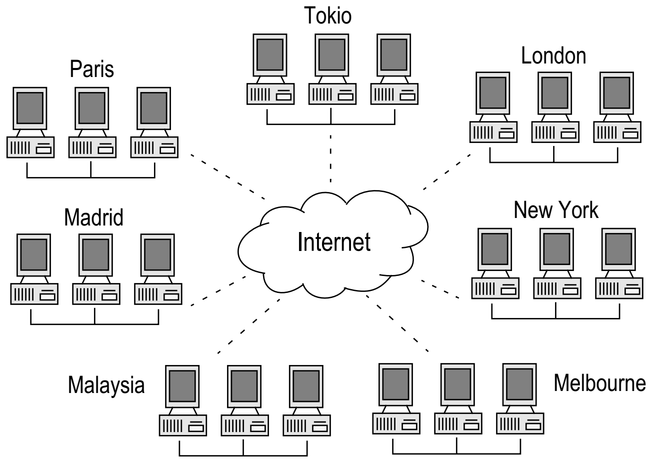


Figure 3: Distribution of Control

the firewall device, utilizing application-specific integrated circuit components for each user connected. In a manner of speaking, we are going to have personalized security from every device that provides us with network service, such as an ATM switch or a communication concentrator.

In any case and, looking at it from hacker's point of view, breaking through system security defences continues to be only question of time: The time necessary to find an error or to have sufficient CPU power that would allow us to discover or to break a specific password. As always, on the security system side, the solution will be to try to increase the time required as long as possible and to increase the CPU power as much as possible. Neither do I believe that we should be excessively pessimistic when it comes to the topic of network privacy: Security is a continuous process and a phenomenon that we should learn to live with in order to try to improve our performance, day by day.

(English by Michael Hird and Richard Butchart)

References

RSA Security.
http://www.rsasecurity.com
Internet Engineering Task Force.
http://www.ietf.org
Secretaría General de Comunicaciones.
http://www.sgc.mfom.es
Agencia Española de Protección de Datos.
https://www.agenciaprotecciondatos.org
EUR-LEX. El Derecho de la Unión Europea.
http://europa.eu.int/eur-lex/es/index.html
Secretariat on Electronic Commerce.U. S. Department of Commerce
http://www.ecommerce.gov

to Contents, page 1

Cookies, Profiles, IP Addresses: Pending Issues in Data Protection Legislation

Rosa Julià Barceló

This article discusses the hottest current privacy issues in the European Union. In particular, the author examines the extent to which European privacy rules apply to US-based companies that gather private data of Europeans through their web sites, taking into account if they use cookies. The article also looks into the issue of the definition of personal data and addresses the question of whether, under EU law, IP addresses are considered personal data and it describes the forthcoming spamming legal regime.

Keywords: European law, personal data, user profiles, cookies, spamming

Rosa J. Barcelò works as a lawyer in the Brussels office of Morrison & Foerster where she specializes in electronic commerce, data protection and intellectual property. Prior to joining Morrison & Foerster, Dr. Barcelò worked as an Assistant Professor of Law at the University of the Balearic Islands and as a Researcher at the University of Namur, where she advised the European Commission on electronic commerce and legal aspects of the Internet. Dr. Barcelò has written numerous articles on electronic commerce and is a frequent lecturer on Internet-related topics. <rbarcelo@mofo.com>

1 Data protection issues in European Legislation

Despite being one of the most complete and exhaustive regulations that exist today concerning the protection of data, the European data protection legislation does not provide answers to some of the new issues that have arisen as a result of the use of the Internet and associated technologies. The same can also be said of the Spanish laws that implements this legislation.

In fact, the two European Directives concerning data protection that were approved at the time the Internet began to be used – the Directive on the protection of the individuals with regard to the processing of personal data and on the free movement of such Data[1], and the Directive concerning the processing of personal data and the protection of privacy in the telecommunications sector[2] – are not very useful to the lawyer, legal professional or academic when they are faced with questions such as:

Should the use of *cookies* be considered to involve the processing of personal data?

Should a cyber navigator's list of visited URLs[3] be classified as personal data?

What requirements should be fulfilled in order to send advertising to WAP telephones?

Should the European Directive regarding the protection of data apply to the processing of personal data conducted abroad with respect to European citizens?

If the data protection laws are short on answers, neither does the competent authority always prove to be a great help. Very often, consulting the officials of the European Commission, the national Data Protection Agencies, or members of Group 29[4] result in either contradictory answers or simply silence. This does not encourage companies with Internet projects to invest in e-commerce, which ultimately damages the development of electronic commerce in Europe.

To deal with some of these problems, the European Commission published a draft proposal on 27 April 2000 relating to the Data Protection Directive, which, among other things, adapts the Directive concerning the processing of personal data and the protection of privacy in the electronic communications sector[5] to the environment of the Internet. At the same time, one must also recognise that the opinions of Group 29 could be of great help, where it not for the fact that their opinions do not always coincide with those of the national Data Protection Agencies. It appears that Group 29 is working on an opinion paper that will provide answers to the questions raised above.

1. Directive 95/46/CE of 24 October 1995 on the protection of individuals with regard to the processing of personal data and the free movement of such data, DOCE No. L 281/31, 23.11.95. This Directive has been transposed into Spanish law through Basic Law 15/1999 of 13 December, regarding the protection of personal data, BOE No. 298, 14.2.99.
2. Directive 97/66/CE of 15 December 1997, concerning the processing of personal data and the protection of privacy in the telecommunication sector, DOCE No. L/24/1, 30.1.98. This Directive has been transposed into Spanish law through Royal Decree 1736/1998 of 5 May that approves the regulation of the Development of Title III of the general Law of telecommunications, BOE No. 213, 5.9.98
3. URLs are the server addresses where Web sites can be found.
4. Group 29: A working party created by Directive 95/46/EC of the European Parliament and the Council of 24 October 1995, which is responsible, among other things, for issuing opinions on the meaning of the Directive.
5. A draft of this Directive can be found at: www.ispo.cec.be/infosoc/telecompolicy/review99/wdprot.pdf

However, at the time of writing, this paper was still confidential, and it will be welcomed when it does appear.

2 Scope of application of the "General" Data Protection Directive[6]

These days, it is very common for Europeans to access the Web pages of ".com" businesses situated in a country outside the European Union, especially the United States, in order to make purchases, participate in bulletin boards, download information, etc[7]. It is usually the case that, in order to carry out these operations, certain personal data must be entered such as, for example, name, address, e-mail address, credit card number, and sometimes even musical preferences, hobbies, etc. One issue that is raised with regard to the processing of this data is whether the rules of the Directive concerning the protection of personal data apply. The issue must also be raised as to what should happen if the opposite situation occurs: that is for example, when a Spanish business established in Spain processes the personal data of citizens of Argentina.

2. 1 Application of the Directive to businesses outside of the European Union

One of the most serious problems of the Directive, as well as its transposition into Spanish law, is the ambiguity with which Article 4 is drafted.

In fact, the scope of application of the Directive is laid down by Article 4 and in particular, Article 4.2 states that the Directive applies in the case where the party responsible for the data processing is not established in a territory of the European Union but uses equipment for the processing of personal data, automated or not, situated in the territory of a Member State, except in the case where the equipment is used only for the purpose of transit through the territory of the European Union. Decree 20 of the Directive complements Article 4 in having a similar criterion: *"Whereas the fact that the processing of data is carried out by a person established in a third country must not stand in the way of the protection of individuals provided for in this Directive; whereas in these cases the processing should be governed by the law of the Member State in which the means used are located and there should be guarantees to ensure that the rights and obligations provided for in this Directive are respected in practice"*. According to Article 4.2 of the Directive, the first obligations that derive from the application of the Directive, by applying the criterion laid down above, are to appoint a representative established in the territory of the State in which the technical means are used and, in the majority of the Member States, to register with the national Data Protection Agency as a data processor.

If we reconsider the case presented, of a business established in a country outside the European Union that collects the personal data of European citizens, for example when they subscribe to an e-mail list, and we try to apply the above rule, several issues arise. First, should the simple "collection of data" through a Web site be considered as "data processing"? Second, can it be considered that the business uses "technical means" located in a Member State of the European Union simply because this company has a Web site that is available to citizens of such Member State?

In answer to the first question, it seems the processing of personal data exists since, as per the Directive, the processing of personal data includes *"the collection, recording, organisation, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction…"*. Therefore, there is no doubt that the simple *collection* of data, regardless of the destination or subsequent purpose, should be considered a "processing of personal data." And this leads us to the second question: can it be considered that *"technical means located in a Member State"* are used when data are collected through Web sites?

As the following examples illustrate, by applying technical criteria, it would be possible to maintain that a US Web site that collects data from European users is indeed using technical means in the European Union. Indeed, if an Internet user accesses a Web site and provides personal information to that site, "technical means" for said processing are normally used in the country where the Internet user is located, which would determine the application of the Directive. These "means" could be *cache* copies or mirror copies, this is, copies of the original content that are copied in the memory cache of servers that are closer to the end user[8]. In other words, a Web site initially hosted in the United States would probably be copied on numerous European servers, thus supporting the view that "technical means" are used in EU Member States. What's more, this is reinforced if the content provider of the Web site utilises *cookies* to collect the personal data. This is because it can be argued that the cookies, kept in the memory of the end user's computer, and therefore in a European country, can be considered as "technical means of collecting data"[9]. As such, it is still important to keep in mind that one could argue that the technical equipment is used only for purposes of transit through the territory of the EU, especially as regards the copies stored in European servers.

Until now, and due to lack of jurisprudence on this issue, these questions remain unanswered or have contradictory answers, depending on the authority consulted. Nevertheless, in recent months, it seems that opinions on the processing of the data of European citizens by businesses located outside of the European Union through Web sites have become stronger among those responsible in the European Commission and Group 29. Unfortunately, none of these opinions can yet be

6. The "general" Directive on the protection of data is Directive 95/46/CE of 24 October 1995, in contrast to the Directive concerning the protection of data in the telecommunications sector, which is directive 97/66/CE of 15 December 1997.
7. Often, these e-businesses have been adapted for citizens of European countries, in that the contents are translated, for example, into Spanish and the prices are published in pesetas.
8. In this way the end user does not have to access the originating site, located further away, which implies greater access time and greater information download time.
9. See Section 3 for a more detailed explanation of how cookies function.

found in an official document. It is hoped that Group 29's forthcoming opinion will provide a comprehensible and balanced solution to this problem.

2. 2 Application of the Directive to the processing of data in third party countries performed by businesses located in Europe

According to article 4.1 (a), the Directive will apply to the processing of personal data where the processing is carried out in the context of the activities of an establishment of the controller in the territory of the Member State. Therefore, for example, if a business located in Spain collects personal data on its Web site, this processing would be covered by the law, regardless of whether the data belongs to European citizens or citizens of other countries. If the processing were carried out by a branch office of this Spanish business located in Argentina, and the data processor responsible for such processing was established in Argentina, then the Directive would not be applicable to this processing, unless the data was transmitted to Spain in some manner.

3 IP addresses, cookies and profiles

Personal data is defined by Section 2 (a) of the Directive as "*any and all information regarding an identified or identifiable individual*". The Directive adds that every person whose identity can be determined, directly or indirectly, shall be considered identifiable; for example, the name of a person identifies him/her directly. Information that identifies an individual indirectly is any and all information that can reasonably be associated to a specific person, such as an identification number, specific physiological or physical identity, psychological, economic, cultural or social elements or characteristics[10].

When the attempt is made to apply this definition to the world of the Internet, numerous doubts arise about certain data that are regularly collected by Web sites as to whether the data should be classified as personal data or not. This is critical because, if they are effectively personal data, the explicit consent of the interested party or owner of the data is needed before this data may be collected and processed[11].

The case of IP addresses[12]

Each computer that is connected to the Internet is identified by a unique number. This address serves to locate the computers that are connected to the Internet. For example, in order for a Web site to be able to send the information that the user wants to download, it needs to have the user's IP address, which tells the Web site operator where the machine is located (for example, the country where the machine is connected and the Internet service provider that connects it to the Internet). This type of IP address is called static, versus dynamic, which is granted by the access provider to the user for specific sessions. The static addresses only disclose the identity of the access provider, but do not disclose any information with regard to the computer of the web surfer.

Can these addresses directly or indirectly be used to identify a person?

Clearly, neither of the two address types identifies a person in a direct manner, even less so a dynamic IP address. Therefore, "in and of themselves" said addresses do not reflect the identity of the user to which they refer. Nevertheless, it is uncertain whether or not they can indirectly identify a Web surfer. It seems that it depends on the use that a person makes of the Web site that collects them. For example, if in some way these addresses are put together or combined with other information obtained in another way (for example the user voluntarily fills in information with his/her e-mail address), there is no doubt that he/she could potentially be identified[13].

Neither of these issues has been an object of jurisprudence, nor is there a decision from Group 29 regarding the matter. However, keeping in mind that it is technically possible for an e-commerce business owner who collects IP addresses to identify indirectly the person associated with an IP address, especially static IP addresses, it seems reasonable for these to be considered personal data. Consequently, and following the terms of the Directive, someone who collects IP addresses must inform the concerned subjects about this and gather their consent before doing so.

Cookies

The same issue is raised regarding the use of cookies. When an Internet user accesses a site on the Web, it is very common for a cookie (that is, a document that includes an identification number) to be sent from this site to the user's hard disk. When the user reconnects to the same site, the site reads the number found in the cookie and checks that the user is already identified and is the same as previously. Probably, in both cases, the visited Web site's program will have recorded each one of the surfer's movements; that is to say, the information that interests the navigator, the most frequently accessed hyperlinks, how much time is spent on each page, and generally his/her preferences. In this manner, whoever sent the cookie can create a profile of the site users, which would allow the user to be offered those items most closely corresponding to his/her preferences on successive visits to the same Web site as identified by the cookie. Thus, for example, upon accessing a Web site dedicated to music sales where user X always chooses classical music, classical music offers will be made to that user. Similarly, a cookie could also tally that user's exposure to a specific announcement, so as to avoid offering the same user the same advertisement on subsequent visits to the page. Finally, besides recognizing the identification number, the cookie can provide additional information regarding the user to the cookie's originator, such as the browser that was used to access a particular

10. The Spanish Law provides the same definition in Section 3 (a).
11. Article 7 of Directive 95/46/CE of 24 October 1995, with regard to the protection of individuals in relation to the processing of personal data and on the free circulation of this data, and Article 6 of Basic Law 15/1999 of 13 December regarding the protection of data of a personal nature.
12. Internet access protocol.
13. Louveaux, S. ECLIP Report on Data Protection (unpublished).

Web site, the language of the program, the user's operating system, e-mail address, etc.[14]

As regards cookies, we reach the same conclusion as with IP addresses: although cookies only identify a computer, they can indirectly identify the person using the computer, especially if linked to supplementary information, and as such should be considered personal data. For that reason, an e-commerce business that sends cookies should inform its users and gather their consent. Otherwise, the processing of the data that is carried out may be considered illegal.

Although Group 29 has yet to develop an opinion, it seems that both Group 29 and the national Data Protection Agencies are inclined towards this interpretation.

4 Spamming

On one hand, many of the new businesses on the Internet see spamming, or the distribution of unsolicited commercial communication, as an important source of publicity and consequently, of earnings. On the other hand, consumer and user associations fight to set up limits on these practices, which they consider an intrusion into an individual's sphere of privacy. The current regulation on spamming in Spain is the result of the transposition of several European Directives.

In Spain, the distribution of commercial communications is regulated by Royal Decree 1736/1998, of 5 May. As per Article 68, unsolicited direct sales telephone calls that are carried out by means of an automatic system (and it is understood that this includes those carried out on devices without human intervention such as faxes) can only be carried out if the user previously consented to receiving them (opt in). As per the same Article, the commercial communication carried out using devices not including those previously mentioned ca be carried out unless the user has manifested his/her desire not to receive then (opt out)[15]. Finally, with regard to this last item, one must keep in mind that the recently approved Directive on electronic commerce establishes the obligation for those who carry out unsolicited commercial communication to consult voluntary exclusions lists; that is, lists of users who have expressed their desire not to receive commercial communications.

As may be observed, this regulation is not technologically neutral, which causes the legal prosecution to differ according to the technology that was used, without having to understand the motive. Thus, unsolicited commercial communications via fax or automatic telephone calls can only be carried out if express consent has been obtained from the user, while commercial communications carried out by other means, such as e-mail (known as spamming) can be carried out unless the user has manifested his/her expressed opposition.

Additionally, technological advances generate doubts regarding legality. For example, it is uncertain if the electronic distribution of messages to WAP telephones, considered a great opportunity for marketing companies, should be governed by one law or the other. While some countries have adopted definitions of automatic telephone calls without human intervention that would cover the inclusion of such advertising messages to WAP telephones, in other countries, such as in Spain, the doubt still remains[16].

To put an end to this uncertainty, the European Commission has proposed an amendment to Directive 97/66/CE of 15 December 1997, concerning the processing of personal data and the protection of privacy in the communications sector. The intention of this draft directive, which is still in the discussion phase, is to submit the distribution of advertising, by whatever means used, to the principle of prior consent from the consumer, referred to as "opt in consent". In other words, the distribution of advertising, whether by means of an e-mail sent to a computer, or messages sent to a WAP telephone, will require of the prior consent of the addressee. Nevertheless, this solution is far from being definitive. On the contrary, the pressure groups that represent the different interests concerned are busily campaining to have their points of view noted. On the one hand, direct marketing businesses state that the development of the Internet in Europe requires that no obstacles be placed in the way of the use of marketing tools, such as e-mail, which favour electronic commerce. On the other hand, data protection agencies, Group 29 and those worried about personal privacy wish to set limits on what are perceived as attacks on personal privacy. Without a doubt, the European Commission, which has declared its desire to maximise promotion of electronic commerce in Europe, is facing a difficult situation[17].

(English by Hilary Green and Rodney Fennemore)

14. Gauthronet, S.; Nathan, F., *On-line services and data protection and the protection of privacy,* Study for the Commission of the European Community, available at http://europa.eu.int/comm/internal_market/en/media/dataprot/links.htm
15. This article implements into Spanish law Article 12 of Directive 97/66/CE of 15 December 1997 regarding the processing of personal data and the protection of privacy in the telecommunications sector.
16. For example, in France through Délibération 85-79, the French Data Protection Agency (Commission Nationale de l'Informatique et des Libertés) considers telephone calls and the distribution of messages to WAP telephones as automatic systems.
17. A description of "eEurope", the European initiative for the information society, may be found at http://www.mcyt.es/infoindustrias/noticias/eEurope.htm.

Some Issues Regarding Encryption Used on the Internet

Pino Caballero Gil

This report contains brief comments on some issues regarding encryption such as the requirement for it, existing risks in the protection of electronic information, and some of the more basic concepts. Special attention is paid to encryption protocols and to legislation in terms of security. Lastly, a conclusion offers a general view of the present and future regarding the problem of security on networks.

Keywords: cryptography, security, encryption, protocols

1 Need

Imagine that you are using the Internet to submit your income tax return, have an extramarital affair or criticize your boss. In any of those cases, you would not want your private e-mail and your confidential documents to be read by a third party, and there is nothing wrong with that. The democratic right to privacy is protected by the Universal Declaration of Human Rights and by the Spanish Constitution. Fortunately, encryption provides the technical tools necessary to uphold this right to privacy. Likewise, if you are carrying out a commercial operation on the Internet, you will face a wide variety of security problems. For example you, as a client, would like to be sure that a virtual shop really exists, that your credit card number remains secret after it is transmitted, and that the information regarding your order one does not change in transit. The merchant also needs to be sure that the information you have sent is authentic. Again, these and many other problems can be resolved by using encryption.

In general, many practical situations arise in which encryption can resolve problems in electronic communication in the public sector (within public administration and with individuals, as is the case with the previously mentioned income tax return and legal proceedings), as in the private sector (among merchants, or between merchants and clients). In e-commerce developed on open networks such as the Internet, some important risks arise that can be resolved using encryption tools. For example, it is important to prevent the situation where the real author of a message could be impersonated by a false author, and also the situation where the message could be altered during transmission or read by an unauthorized person. There is also a significant risk that the transmitter of a message could deny having transmitted it or that the addressee could deny having received it. Without solutions to these problems, there would be no guarantee regarding the author of an electronic message, nor its content, nor even its very existence, which would inevitably lead to the invalidity of electronic commercial transactions.

2 Concepts

In the world of cryptology, there are two types of people. On one hand, there are the users who want to keep their communications secure (encryption), and on the other hand there are the infiltrators who try to intercept these messages (cryptanalysis). For the former, the objective is that the infiltrator should understand as little as possible of his messages, whilst the infiltrator himself wants to understand them all easily. Cryptology is a continuous battle between these two groups. An infiltrator's success results in the need to improve cryptographic methods, which implies a new challenge for the infiltrator, and so on and so on. This means that cryptology is a living science, since as soon as details of a new encryption method are announced, the possibility of attacking it arises automatically. Clearly, those who are proposing the new method are aware of this, and it is common practice for possible ways of breaking the code to be enumerated alongside its description, along with reasons why these approaches would probably fail. Of course this is no indication of the likelihood of a different approach succeeding.

Before the 70s, encryption was too complicated and expensive for everyday use. Then two inventions opened up new horizons. Firstly, two U.S. mathematicians from Stanford University, Diffie and Hellman, thought up a new encryption method, public key encryption, based on a large number of calculations. Secondly, new, quick and affordable microcomputers gave users the ability to use this public key encryption. Now, barely twenty years later, encryption is within everyone's reach, and its use has become commonplace.

But why was the introduction of public key encryption so crucial? In traditional systems, i.e. those with private keys, the same key is used to encode as well as to decode, which implies the need for the transmitter and receiver to have previously communicated the key to be used, by means of a secure chan-

Dr. Pino Caballero Gil (Gran Canaria, 1968) is a professor at the University of La Laguna. She received her Ph.D. in Mathematics in 1995 with a thesis on "Code in Flow." In 1996, she published one of the first handbooks in the Spanish language on "Introduction to Encryption." She currently presides over the Encryption group of the Department of Statistics, Operative and Computational Research.<fcaballe@ull.es>

to Contents, page 1

nel. Public key encryption, or asymmetric encryption, eliminates this problem, as previous communication and a secure channel are unnecessary, and additionally, the number of keys each user has to handle decreases proportionally to the number of people he communicates with. This type of encryption is based on the use of a pair of associated keys: a private one used to decode, and a public one used to encode. The public key encryption method most widely used is RSA, based on the problem of factorizing the product of two large prime numbers. This system is implemented in the majority of security protocols.

However, it has to be said that public key encryption also has a disadvantage, which is that it involves a large number of mathematical calculations, which is the reason for its relative slowness compared with the speed of private key encryption. In fact, for very large messages, or when rapid communication is crucial, private key encryption is generally used today, although technological advances have minimized the difference in speed between the two techniques. A widely-used trick to take advantage of the public key encryption concept without jeopardizing the efficiency of the communication is based on the use of a hybrid system which involves using symmetric encryption to encode the message, adding the private key at the end, coded using public key encryption.

Public key encryption is also used for the digital signature of messages, since it enables the identity of the sender to be verified easily. When Benito (B) sends a message to Alicia (A), he first encodes it with his private key and then re-encodes it, this time using A's public key. When A receives it, she first decodes it using her private key, and then with B's public key. If the end result makes sense, A can be sure that the transmitter was B, since the only one who has the decoding key is B. When a public key is used to digitally sign a very lengthy message, the slowness hurdle can easily be overcome by applying a hash function that transforms the message into a shorter one. This function is characterized by its unidirectional nature, since the original message cannot be derived from the end result, and also by the fact that it is computationally impossible to find another message with the same result.

In digital signatures based on public keys, the following problem arises. When B digitally signs a message and sends it to A, how can A be certain that the B's public key really belongs to B? An impostor could have infiltrated the message and pretended to be B, sending A a false public key that allows the use of a corresponding false private key to decode messages that A unsuspectingly sends encoded with the supposed public key of B.

So a secure method of distributing public keys is needed. To date, two possible alternative methods have been proposed. The one used in the PGP system is based on the principle of "trust the friends of my friends", i.e. that public keys be verified by trustworthy people. This system has the advantage of not being centralized, but on the other hand, it makes secure communications with people with whom one does not share associates impossible, and additionally, in the case of a dispute, it is not clear where the responsibility lies.

The other technical solution for the distribution of public keys is based on the concept of a Certification Authority, whose function is, as its name indicates, to issue digital certificates that guarantee that each public key belongs to its legitimate owner. This solution is the one that has proved more popular, since it resolves the two problems mentioned above. The digital certificates issued by this authority provide Internet users, organizations and businesses with simple methods of mutually checking identities. For the benefit of the client, the certificates constitute:

- A simple method of verifying the authenticity of a business before sending confidential information.
- A guarantee that, in the worst case, users can obtain the physical address and the legal name of the business in order to initiate legal action against it.

For the merchant, the digital certificates provide:

- A simple method of verifying the e-mail address and the identity of an individual.

Nevertheless, it must be emphasized that authentication by means of digital certificates does not prove that a user is who he claims to be; since it only shows that he possesses a specific private key signed by the Certification Authority. Even so, to date, digital certificates constitute the most secure method of identifying users on the Internet.

3 Protocols

In the previous section, we have mentioned the two most commonly known aspects of encryption, which are the protection of confidentiality, using codes, and the protection of authenticity, using digital signatures. This section looks at a variety of cryptographic techniques developed to resolve a series of problems that go beyond those of authentication and privacy. They relate to algorithms based on codes and are known as cryptographic protocols.

Many everyday situations can be resolved in the world of telecommunications by means of an adequate cryptographic protocol. Examples of this are the launch of currencies, the finalization of contracts and voting. E-commerce is a separate case, in which diverse problems are resolvable with cryptographic protocols. Thus in almost all of the more well known encryption symposiums, such as the U.S. Crypto, the European Euro crypt and the Asian Asia crypt, a new problem is addressed along with a new cryptographic protocol to resolve it.

A familiar situation is the existence of a secret so important that is it advisable not to leave it in a single pair of hands, for example the key to access a high security system. In this case, the best solution consists of dividing the secret into several parts and distributing it among various users, so that none of them knows the secret in its totality, but if they get together, they can reconstruct it. The type of cryptographic protocol developed to resolve this situation is known as the threshold plan or the sharing of secrets, and various proposals exist based on polynomials, vectors, prime numbers and matrices, among others.

Two users wish to exchange secret information by means of a third party that they both trust. The idea of the protocol known as subliminal channel consists in transmitting information that appears innocent, but that subliminally hides another message. For example, A and B could simply previously agree that an odd number of words signifies a 1, and an even number implies a 0, and use this agreement to transmit secret messages in an open conversation.

A previously mentioned problem is the realization on the part of two users of an experiment with two possible, equally probable results, such as the launch of a currency, but where the two launchers are so distant from each other that they have no way of checking whether the other is cheating or not. The corresponding protocol receives the name of the launch of currencies. Similar to this example is that of the protocol developed to play poker when the players are not physically in the same place, from which we have the protocol that allows the assurance of clean play in mental poker.

Let us now centre our attention on one of the most fascinating problems. Let's suppose that A (the tier) knows some verifiable information, such as a private key, and that A wishes to convince B (the verifier), beyond reasonable doubt, that A possesses this information. A could simply show B the information, but this has the disadvantage that B would then know the information and could show it to a third party, pretending to be A. In the protocol known as demonstration of null knowledge, A convinces B that she has the information, but does not disclose a single bit. Consequently, B is fairly sure that A has the information, but does not possess the information itself, and could not pretend to a third party that he was A. Proposals for demonstrations of null knowledge exist based on tools as different as numerical theories, graphs or boolean logic.

A protocol known as unconscious transfer allows the following problem to be resolved. User A wishes to transfer a secret to user B, so that when B receives it, A does not know if it has actually arrived or not. This type of protocol is one of the basic tools or more advanced protocols that resolve the problem of the finalization of contracts when the parties are not in the same physical location. In these, a gradual exchange of information takes place that is fundamental to solving the problem of non-acknowledgement of receipt at destination.

Lastly, very useful cryptographic protocols have been designed for use in systems that allow voting by e-mail, ensuring that voters are authentic, that their votes are unique and secret and are included in the final count.

4 Legislation

In 1998 a new General Law of Telecommunications was approved, that states: "… the obligation of notifying a body of the General Administration of the State or a public agency, regarding the algorithms or any coding procedure utilized may be imposed... Network or telecommunications services operators that use any coding procedure must facilitate to the General Administration of the State, without cost to this body and with the purpose of an opportune inspection, the decoding devices used…". It is clear that this law was written for the benefit of authorities in their fight against crime and terrorism.

However, encryption currently constitutes the security base of e-commerce and of the general user computer network, with its two most common uses being the protection of e-mail messages and the secure transmission of credit card numbers between individuals and businesses or banks. In fact, the non-governmental demand for encryption grows each day, and the above law, which could result in the problematical establishment of a compulsory system of centralized storage of keys, presents a conflict between individual privacy and the public interest. Those hypothetical centralized databases would be a very tempting objective for delinquents, and the volume of e-commerce would decrease due to clients' fears that their credit card numbers could fall in the hands of third parties. However, a criminal would have no difficulty in carrying out his activity because, in practice, it would not be too difficult for him to develop cryptographic software of his own, into which ordinary users would enter their keys. Additionally, since the reason for the existence of encryption is to make private communication truly private, if a third party could read this private communication, there would no longer be any sense in using encryption.

On the other hand, everyone is talking about legislation with respect to digital signatures and e-commerce. In several European countries and states of the U.S., laws have been approved or are under review regarding e-commerce and e-signatures. For Europe, in 1997, two documents were presented in which several measures were to take effect before the year 2000 to promote e-commerce and to generate trust in digital signatures and in encryption. These measures related to the implementation of secure technologies and to the judicial and institutional frameworks whose aim is to support and avoid contradictions in the legislation of the different member states.

In Spain, in the context of the public administration, a law was approved in 1996 entrusting the National Mint with providing security services with regard to electronic communication, for example the issuing of digital certificates. On the other hand, in September 1999, a Royal Decree was approved regarding electronic signatures that envisaged a system of digital certificates issued by trustworthy third parties and that, in particular, established the registration of certification services dependent on the Department of Justice. In this Decree, digital signatures are considered secure when based on public keys, and as such are binding on the signatory and valid in a court of law.

5 On the Internet

Typically, a web server is considered secure when it carries out certain cryptographic protocols that protect the information transferred between the server and the browser. The encryption included in such protocols is widely recognized as a prerequisite for Internet commerce, and a weak code is typically sufficient for the majority of transactions since it is always easier for an infiltrator to search the Internet for uncoded credit card numbers.

In general, the problem of security on the Internet consists of several fundamental parts:

- Secure the server and the data it contains.
- Secure the information that travels between the server and the user.
- Secure the user's computer.
- Verify the identity of the user to the server.
- Verify the identity of the server to the user.

One of the main issues to resolve on networks is that of access, to the server as well as to the user's computer. Although many identification systems exist which are not based on encryption, for example, biometrics, search engines with counter keys, etc., the majority can be improved by means of the use of digital signatures, and these can in turn be improved using smart cards to generate the pair of private and public keys, and to store the private one.

When information travels across the Internet, the main risk to avoid is unauthorized interception, and the most practical solution for this is encryption. One of Netscape's first innovations was SSL, a system for automatically encoding information as it travels by Internet, and decoding it upon arrival. In fact, in recent years, over a dozen protocols and cryptographic systems for the Internet have been developed. The most widely used to date can be classified in two main categories: the security protocols used on networks, such as SSL, and the ones that are used to encode e-mail messages, for example PGP. In both cases, confidentiality, authenticity of transmitter, integrity of message and non-denial of origin is provided.

6 Future

You don't need a crystal ball to predict that encryption will experience an even quicker growth in the future than it has experienced up to now. Some possible new applications:

- Since many telephone calls are processed via satellite, it is logical that important calls be coded.
- Authentication plans mean that people who try to evade the cost of pay television will not succeed.
- In multi-user computer systems, each user will conveniently be identified by means of smart cards that will provide a greater degree of security than keys currently being used.
- With the growing use of electronic banking, the digital signature will become more and more indispensable.
- Data authentication methods will be used as a defence against computer viruses.
- In the development of e-commerce, for small purchases as well as large negotiations, codes, digital signatures and cryptographic protocols will play a major role.

Finally, the encryption of public keys and digital certificates appears to offer the best prospect of sending secure and authentic electronic messages on open networks, thus facilitating e-commerce.

References

[Caballero 96]
P. Caballero: "Introducción a la Criptografía", Ra-Ma, 1996.

[Garfinkel/Spafford 99]
S. Garfinkel, G. Spafford: "Seguridad y Comercio en el Web", McGraw-Hill,1999.

[Martínez Nadal 98]
A. Martínez Nadal: "Comercio Electrónico, Firma Digital y Autoridades de Certificación", Civitas, 1998.

[Schneier 93]
B. Schneier: "Applied Cryptography", Wiley & Sons, 1993.

[Stallings 95]
W. Stallings: "Network and Internetwork Security Principles and Practice", Prentice-Hall, 1995.

Security through Business Process Re-Engineering

Gaby Herrmann, Günther Pernul

Apart from the traditional way of doing business, computer based execution of business is becoming increasingly important. This applies to relationships both with clients and business partners. In relationships security plays an important role [Enquête 98]. In the traditional business processes security is usually satisfied intuitively. With computer-based transactions this is no longer possible. Therefore, re-engineering of business processes is necessary in order to achieve the required security. In this article we give an insight into the problem of the security in business processes and introduce an action model for the implementation of secure business processes. We illustrate the action model proposed with an example focusing on the security requirement "non-repudiation".

Keywords: business processes, security, re-engineering, digital signature, non-repudiation

1 Introduction

Apart from the traditional execution of business processes, computer based business process execution is becoming increasingly important. This applies to relationships with both clients and business partners. In these relationships, security plays an important role. Existing security standards are changing due to the trend from executing business processes without (or only with little) computer assistance to handling these processes extensively using electronic means. This change may effect the meaning of security standards (such as confidentiality of a message, possibly delivered electronically). Or the person responsible for carrying out a business process may become aware of the necessity of security requirements (such as the authentication of a signature).

In the traditional execution of business processes security requirements were usually intuitive and developed through long-term experience with the business practices. Confidentiality of a written message is usually ensured by the notice "confidential" on the envelope, as envelopes usually prevent unauthorized reading of a letter. If the confidentiality of a message to be delivered is so high that a normal delivery of the letter cannot sufficiently guarantee it, then the message can be delivered personally by a person trusted by the sender. These methods have been known and practised for many centuries, so that they are proven, and the people involved are accustomed to them. However, if a business process and also the delivery of a confidential message is done purely electronically, then the people involved have no intuitive knowledge how to conduct such business. Therefore, in order to ensure security, the necessary procedures must be defined in the business process execution models. This requires a re-engineering of the business processes.

In this article we introduce an action model that presents security requirements in business processes and illustrates these procedures with the non-repudiation requirement of an agreement as an example. After discussing the importance of security in business process management we will introduce an action model supporting the implementation of security requirements for business processes. Finally, we will illustrate this with the example of not being able to deny having digitally signed an agreement (non-repudiation).

2 Security and business process management

Currently, there is no general understanding of the term "business process". Different authors use different definitions. For us, *business process* refers to a variety of related activities that support the implementation of corporate goals, [Gausmeier/Fahrwinkel 94], and which aim at efficiently fulfilling a client order [Georgakopoulos et al. 95] (including orders from clients within the company).

If a business process which used to be processed without or only with little computer-aided technology is to be increasingly supported using computers, then this business process needs to be analysed in depth and the appropriate business process model needs to be adapted accordingly. In this procedure, the necessary security requirements have to be taken into account, as

Günther Pernul is professor at the University of Essen. He is head of the study group Information Systems. Before that he worked at the Institute for Applied Information Technology and Information Systems, University of Vienna. He was guest scientist at the Database Systems Research and Development Centre of the University of Florida and at the College of Computing at the Georgia Institute of Technology. His research interests are Electronic Commerce, Security and other aspects of databases. <pernul@wi-inf.uni-essen.de>

Gaby Herrmann studied Information Technology at the University of Karlsruhe (TH). Since 1992, she has been a research assistent at the University of Essen, until 1995 in the study group Communication Systems, later in the study group Information Systems. Her research interests are in the areas of Workflow Management, business process modelling and modelling of security semantics. <herrmann@wi-inf.uni-essen.de>

to Contents, page 1

	Security components	Description method	Support methods
Level 3	abstract specification of business process security requirements	graphic representation method	• graphic concepts for the representation of security requirements • tools to check syntactic and semantic correctness of security requirements *Level 3.2* *Level 3.1* collection of already modelled flows for the implementation of security requirements in the form of reference models
Level 2	detailed specification of basic security elements or of more complex security-relevant processes	ALMO$T specification language	collection of flows for the implementation of basic security elements or more complex security-relevant flows using the basic elements of level 1
Level 1	hardware security elements, software security elements, supporting elements and services	Programs, program modules, hardware	collection of hardware and software basic elements (e.g. crypto-library, APIs, security dongles, e-mail system addresses of providers of security-services)

Fig. 1: Three layer architecture of the action model [Herrmann/Pernul 99]

these standards can no longer be applied intuitively and based on long-term experience, as has been the case until just recently. Therefore, in order to ensure security, the required procedures must be defined in the business process processing models. This requires a re-engineering of the business processes within the business process management. *business process management* is understood as a concept for the formation, coordination, and execution of business processes based on a certain model. [Scheer et al. 95] business process management includes the modelling of business processs, the analysis, modification, and simulation of these transactions to ensure correctness, optimization, and proper execution [Gruhn/Wolf 94].

Security requirements concerning business processes can stem from a variety of causes and can refer to different elements of business processes are listed below. business process elements, which may appear as security objects[1]:

- *Information* in the instance of procedure, end product, and other information
- *Agents* in the instance of executing agent and commissioning/commissioned agent
- *Information flow*
- *Activities*

The security requirements we identified so far can be organized into the following categories:

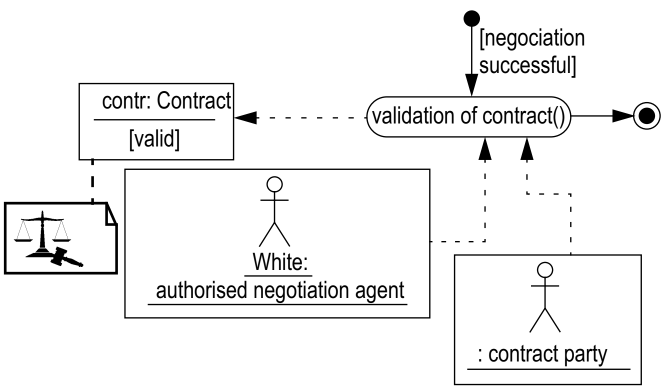


Fig. 2: Perspective of business process (extract)

- *General requirements:* confidentiality, integrity, availability
- *Intellectual/physical property:* originality, authenticity, copyright, right of ownership
- *Relationships:* interdependence, non-repudiation
- *Data protection:* anonymity, pseudonymity, privacy

The security requirement "covering up activities" cannot be incorporated into any of these categories[2]. Furthermore, not each of the itemized security requirements is relevant for each kind of business process element [Herrmann 99].

3 Action model

In support of the re-engineering process concerning security requirements, an action model was developed in the project MoSS (**Mo**delling **S**ecurity **S**emantics); this action model's architecture is illustrated in figure 1.

Before the modification of business processes in response to security requirements can be started, the modifications have to be specified in the corresponding business process model. For this purpose, it is necessary to expand the specification language used by documenting the security requirements. For the representation of business processes, several points of view have to be taken into account [Curtis et al. 92]. MoSS considers the functional, static, dynamic, and the organizational perspectives of a business process and offers a partly comprehensive and/or referential point of view for a better understanding of a business process (business process perspective) [Herrmann/Pernul 99]. MoSS is not restricted to a particular specification language. A commonly used specification language is the **U**nified **M**odelling **L**anguage (UML) [Booch et al. 99].

2. There is a close relationship between the security requirement "covering up of activities" and the security requirement "confidentiality". Since however confidentiality is generally seen only in the context of the existence of information or agents and their qualities, we quote the requirement "covering up of activities" as an independent requirement and not the same category as "confidentiality" (category "general requirements"). Furthermore, the relevance of the categories "confidentiality" and "covering up of activities" doesn't include the same classes of business process elements.

UML lets the user illustrate security requirements through, for example, previously defined stereotypes. As not all security requirements are relevant for each business process element, the user should have support through a tool for the syntactic checking of the assignments (fig. 1, Level 3.2).

Detailed models of the examined business process must also show the individual steps that are necessary for the realization of the specified security requirements. For this, knowledge about the realization of security requirements is necessary. This knowledge may be gained from a security expert, through case studies, or through reference cases. As corporate expectations concerning security requirements and the organizational environment for achieving these requirements may differ from case to case[3], case studies are summarized and the results, which have been gained through the development of these case studies, will be made available to the user in the form of reference models (level 3.1). These reference models are meant to ensure a level of detail for achieving security requirements, using exclusively fundamental security elements or more complex security relevant activities for which realizations on level 2 or 1 already exist. Fundamental security elements exclusively describe actions, which are used for achieving security requirements, for example Authentication (communication partner).

On architectural level 2, there is a collection of descriptions of realizations of fundamental security elements and of more complex, with security requirements equipped actions. These descriptions use the specification language ALMO$T [Röhm et al. 99] and allow the direct conversion of the corresponding fundamental security element and/or the activity, guided by security requirements, with the assistance of existing software and hardware chips (level 1), if the necessary chips exist.

If realization of the required fundamental security elements and/or security relevant activities is impossible, the employees responsible for the business process need to be contacted in order to discuss the possibility of a reduction of security requirements to an extent that would allow the requirements to be attained. If this is impossible, the business transaction cannot be carried out considering the specified security requirements.

4 Non-repudiation of an agreement

An example will explain the above described procedure of the realization of a secure business processes through reengineering. This will be simplified by using case studies instead of reference cases on level 3.1 of the architecture.

An agreement between two parties serves as an example. One possibility to establish an agreement is by creating a written statement, signed by both parties. This is a measure appropriate to the traditional methods of carrying out business. How can the non-repudiation of an agreement[4] be ensured, if the agreement was concluded using exclusively electronic means?

Digital signatures are the technical equivalent to the traditional written signatures. To which degree digital signatures may have legal equivalence to written signatures depends on the prevailing laws. In many countries, corresponding laws are already in effect (e.g. [IuKD 97]), and the EU has developed a guideline concerning common basic guidelines for electronic signatures [EU 99]. However, we are not aware of any final legal decisions concerning this issue. As all legal efforts concerning digital signatures depend on their technical characteristics, the presented modifications of a business process, in which an

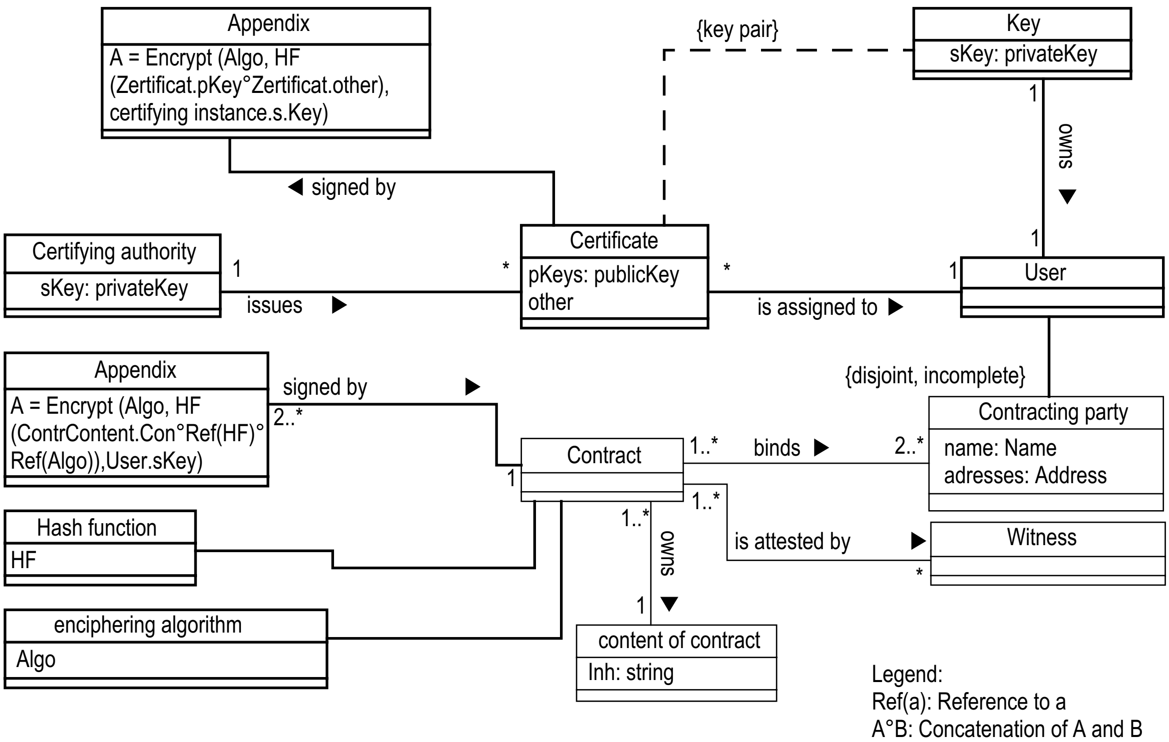


Fig. 3: Static Perspective (extract)

to Contents, page 1

agreement between two parties has been found and for which non-repudiation is requested, are based on the Authentication Framework of the ITU [ITU 97].

The individual perspectives of the examined business processes are illustrated using UML. Only aspects relevant for the non-repudiation of the agreement are shown.

The business process perspective portrays a contractual agreement between the authorized representative White and a representative of the other negotiating party. The agreement (portrayed as the contract) shall be demonstrable.

A digital signature under a text guarantees that exactly this text was signed exactly by the person who own the private keys pertaining to the digital signature. Therefore, a specific person may be assigned to a specific digital signature under a signed text. The Authentication Framework of the ITU requires that a digital signature shall be realized in the form of an appendix. The digital signature is created through the application of an undetermined one-way Hash function on the text to be signed and the encryption of the created Hash value with the private key of the signing person. The signed text must show the applied Hash function and encryption algorithm used. A digital signature may be checked by applying the same Hash function to the signed text and comparing this result with the result of the decryption of the digital signature utilizing the public key of the person signing. According to ITU, a certificate holds a users public key, as well as additional information, and is protected against falsification through encryption by the certifying authority using its private key. The ITU's measures for the use of digital signatures influence the different perspectives of a business process model, if the corresponding business process requires the use of digital signatures.

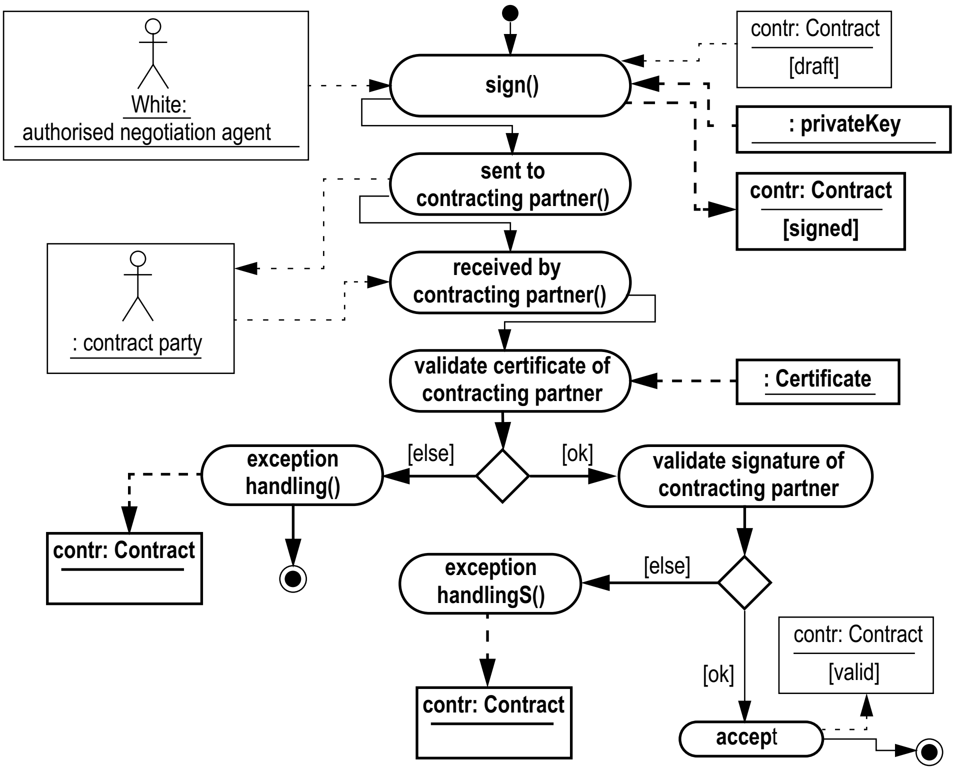


Fig. 4: Functional Perspective (extract)

The static perspective represents the information and products required for the business process. The changes required based on the request for non-repudiation (refer to the example of the business process perspective, fig. 2), are highlighted in bold in Figure 3. The contract is expanded by one appendix per contract party, containing the pertinent digital signature. The signed text contains the contract content, as well as references concerning the applied encryption algorithms and utilized hash functions. Each contract party must own a signed certificate issued by the certifying authority, which certifies the public key pertaining to the used secret key.

As the procedure concerning digital signing of an agreement is not yet known by potential signers, the corresponding steps are explained in the functional perspective of the business process model (fig. 4). The functional perspective explains the applied activities and the flow of data between them.

As there is no guarantee that encryption procedures and keys which are considered secure today will remain secure in future, the EU guideline has determined a period of validity for certificates and for certified keys. If non-repudiation of an agreement is requested for a period longer than the period of validity of the used keys, then further steps will be necessary, including during the period of archiving of the contract. The contract partner whose certificate is approaching the end of the period of validity, for example, may be asked to sign the agreement once again, using a key certified for an extended period of time. If the contract partner refused to comply, it is possible to include a trusted third party to testify the validity of the digital signature of the contract partner. The problem of a limited validity of a certificate shows that security requirements on a business process do not only affect this process, but may also affect other business processes (i.e. the archiving process).

The organizational perspective of a business process model describes the construction of a corporate structure. Fig. 5 shows sections of the organizational perspective of a business requesting the non-repudiation of an electronically concluded agreement. The employees, which may act as authorized negotiating partners, must be able to use the corresponding secret key in order to give their signature. Additionally, the position of signature manager has to be created who will take care of the problem of expiring certificates in the archiving department.

The dynamic perspective of a business process model which describes the life cycle of information objects is subject to

to Contents, page 1

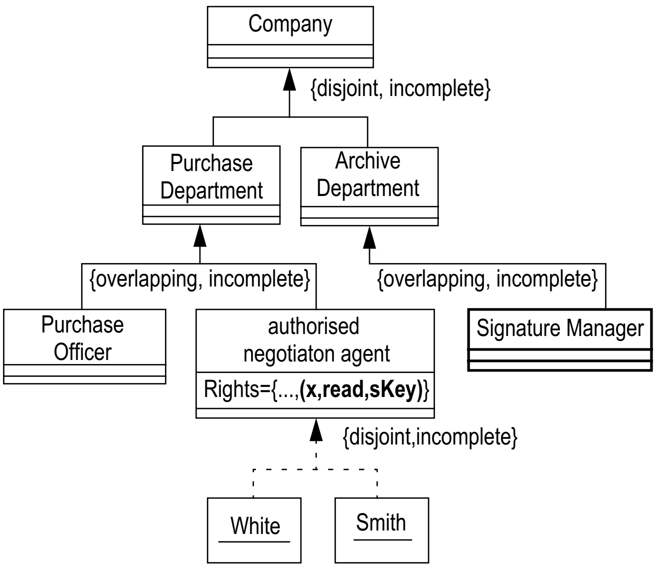


Fig. 5: organisation Perspective (extract)

changes in view of the non-repudiation of an electronically concluded agreement, too. For example, in an agreement concluded in the traditional method, the information object "contract" changes from "draft" directly to "valid", as all contract partners sign approximately at the same time. An electronically concluded draft of a contract which will get electronic signatures, may have a signature process which may have physical and temporal differences. The agreement is not valid during the period between the giving of the signature of all contract partners involved. The agreement of the party which has already signed is already demonstrable, while non-repudiation of the agreement of the other party is not yet given.

The modifications of perspectives of business process models are illustrated in figures 3 to 5 and may be included in the collection of case studies on level 3.1 of the architecture; they may be re-applied in the case of a new request for non-repudiation of an agreement.

5 Conclusion

The consideration of security becomes increasingly important due to the increasing computer-aided execution of business processes. In the traditional execution of business processes, sufficient security standards were usually intuitive and based on long-term experience. This is no longer possible with computer aided transactions. The necessary procedures must be put down in writing in the business process models. As the realization of security requirements of a business process often affects different perspectives of the business process model, and even possible other processes of the business, a comprehensive business process re-engineering is inevitable.

The procedure introduced in this article offers support for this problem. It is a structured procedure which aids in reducing complexity and which supports the re-use of already existing modules. This modules include hardware, software and also descriptions of possible procedures.

References

[Booch et al. 99]
G. Booch, J. Rumbaugh, I. Jacobson: The unified modelling language user guide. Addison-Wesley, 1999.

[Curtis et al. 92]
Curtis, B.; Kellner, M. I.; Over, J.: Process modelling. Communications of the ACM, September 1992. Vol. 35, No. 9, pp. 75–90.

[Enquête 98]
Electronic Commerce Enquête 1997/98. Computerzeitung, Gemini Consulting GmbH and Universität Freiburg, Institut für Informatik und Gesellschaft (IIG). (in German)

[EU 99]
Richtlinie 1999/ /EG des Europäischen Parlaments und des Rates vom über gemeinschaftliche Rahmenbedingungen für elektronische Signaturen, 1999. (in German)
http://www.ispo.cec.be/eif/policy/diresde.doc (26.7.2000)

[Gausmeier/Fahrwinkel 94]
J. Gausmeier, U. Fahrwinkel: Strategiekonforme Geschäftsprozesse und CIM-Massnahmen. CIM-Management, 10 (2) 1994, S. 58–61. (in German)

[Georgakopoulos et al. 95]
D. Georgakopoulos, M. Hornick, A. Sheth: An Overview of Workflow Management: From Process Modelling to Workflow Automation Infrastructure. Distributed and Parallel Databases, 3, 1995, pp. 119–153.

[Gruhn/Wolf 94]
V. Gruhn, S. Wolf: Software-Entwicklung auf der Basis von Geschäftsprozess-Management. In HMD, Jahrgang 31, Heft 180, 1994, S. 116–125. (in German)

[Herrmann/Pernul 99]
G. Herrmann, G. Pernul: Viewing Business Process Security from Different Perspectives. International Journal of Electronic Commerce, Vol. 3, No. 3, 1999, pp. 89–103.

[Herrmann 99]
G. Herrmann: Security and Integrity Requirements of Business Processes - Analysis and Approach to Support their Realisation. Proceedings of CAiSE*99 6th Doctoral Consortium on Advanced Information Systems Engineering, Heidelberg, 14.–15. June, 1999, pp. 36–47.

[ITU 97]
International Telecommunication Union: X.509, Information Technology – Open Systems Interconnection – The Directory: Authentication Framework, 06/97.

[IuKD 97]
Gesetz zur Regelung der Rahmenbedingungen für Informations- und Kommunikationsdienste (Informations- und Kommunikationsdienste-Gesetz – IuKDG) in der Fassung des Beschlusses des Deutschen Bundestages vom 13. Juni 1997 (BT-Drs. 13/7934 vom 11.06.1997). (in German)

[Röhm et al. 99]
A. Röhm, G. Herrmann, G, Pernul: A Language for Modeling Secure Business Transactions. Proceedings of the Fifeteenth Annual Computer Security Application Conference ACSAC, 1999, pp. 22–31.

[Scheer et al. 95]
A.-W. Scheer, M. Nüttgens, V. Zimmermann: Rahmenkonzept für ein integriertes Geschäftsprozeßmanagement. In Wirtschaftsinformatik, Nr. 5, 37. Jahrgang, 1995, S. 426–434. (in German)

[Starke 94]
G. Starke: Business Models and their Description. In Chroust, G.; Benczur, A. (Hrsg.): Workflow Management: challenges, paradigms and products. CON'94. Schriftenreihe der Österreichischen Computer Gesellschaft, Band 76, Oldenbourg Wien München, 1994; S. 134–147.

to Contents, page 1

A Model for Security in Agent-based Workflows

Henrik Stormer, Konstantin Knorr, Jan H.P. Eloff

With the rise of global networks like the Internet the importance of workflow environments is growing. However, security questions in such environments often only address secure communication. Other important topics are role-based access control and separation of duty. This paper shows how mobile agents can be used to implement these and other security features in workflows.

Keywords: agent, fraud, role, security, separation of duties, workflow

1 Introduction

Workflow environments have hugely benefited from the technical advancements made available by the Internet over the last years. Many workflow environments today are implemented over public networks such as the Internet. Because workflow environments in most situations represent the "bread-and-butter" of a company, the implementation thereof has raised serious information security problems. Organizations are concerned about their privacy on the net as well as of the privacy of client information. Similar to other systems the information security requirements of a workflow system are modelled on the ISO 7498-2 standard. This standard proposes the following information security services: identification and authentication, authorization (access control), confidentiality, integrity, and non-repudiation. Mechanisms for each of these services must be employed to secure a workflow environment.

Identification and authentication, confidentiality and non-repudiation services are implemented similarly to those in non-workflow environments. The services of authorizatiyon (with the main focus on access control) and integrity require special design considerations and implementation details. For example access control requires the modelling of access based on the type of tasks to be performed on the objects travelling around in a workflow environment. A unique feature of integrity in a workflow environment is to preserve the contents of objects according to business rules. These business rules are linked to the operational characteristics of an organization. There is a need for new approaches modelling the design and implementation of the access control and integrity services in workflow environments.

Currently available research results in the area of access control are dominated by models of role-based access control (RBAC). RBAC shows good potential to be successfully employed in a workflow system. The information security principle of separation of duties (SoD) is important in the modelling of integrity in a workflow environment. A physical and logical separation of tasks can improve the prevention of fraudulent activities.

Agent technology shows great potential in the field of workflow systems. Furthermore, information security aspects like RBAC and SoD can be considered in the agent-based implementation of workflows.

Therefore, the primary aim of this paper is to give an architectural model and a framework for implementing access control and integrity in a workflow environment. Intelligent and mobile agents are applied to current workflow technology to meet access control and integrity requirements.

The remainder of the paper has the following structure: Section 2 gives an introduction to workflow environments and agents. Section 3 describes a sample process which will be used for illustration purposes throughout the paper. An architecture for agent-based workflows and its different agent types are discussed in Section 4. Section 5 introduces the notions of RBAC and SoD and gives a formal model for them within the workflow environment. Implementation of several security features in an agent-based workflow system is the topic of Section 6. Section 7 gives a conclusion.

Henrik Stormer has studied computer science at the University of Saarbrücken (Germany). He is currently a PhD student at the Department of Information Technology at the University of Zurich. His research interests are mobile agents and agent-based workflow systems. <stormer@ifi.unizh.ch>

Konstantin Knorr studied mathematics at the universities of Mainz and Frankfurt (Germany) and is currently a PhD student at the Department of Information Technology, University of Zurich. His research interest are formal models of security and security in workflow environments. <knorr@ifi.unizh.ch>

Jan Eloff is a professor in computer science at Rand Afrikaans University, South Africa. He is currently a visiting professor at the University of Zurich, Switzerland. He is chairperson of the Special Interest Group in Information Security and is chairperson of the International Working Group 11.2 of IFIP specialising in small systems security. He is an evaluated researcher from the National Science Foundation, South Africa. <eloff@rkw.rau.ac.za>

to Contents, page 1

2 Background

This section gives background information on workflow management issues and mobile agents.

Workflow management is an essential research area in computer science. A workflow is an executable business process whose modelling and execution is supported by a software system called workflow management system (WfMS) [Georgakopoulos et al. 95]. Before a workflow can be executed, it has to be described in a way the WfMS is able to understand. This description is called a *process definition*. The definition has to be made during *build time* before a workflow can be executed. During run time of the system many instances of the workflow are generated according to the process definition. The main elements of a process definition are tasks, objects, subjects, roles, and the control flow. A process consists of several tasks whose chronological and logical order is given through the control flow. To describe a task, it has to be specified which roles are granted access to which objects. Subjects can be associated with persons but also with machines and computer programs [Leymann/Altenhuber 94].

The *Object Management Group* defines a software agent as "a computer program that acts autonomously on behalf of a person or organization" [Crystaliz et al. 97]. The following properties characterize agents:

- pro-active (support of the user's work)
- adaptive (learning the user's preferences or the ability to work on different platforms)
- autonomous (limited communication with its creator)
- intelligent (making 'intelligent' decisions [Ferber 99])
- mobile (can actively migrate in networks to different systems and move directly to the local resources, like databases or application servers)

Before agents can be used, each system needs to install a so called *agent-place* to create, delete and execute agents. Agents can migrate from an agent place to another performing the work locally. Lange and Oshima [Lange/Oshima 99] give reasons why to use agents: e.g. reduction of the network load, overcoming of network latency and encapsulation of protocols.

An agent-based workflow is a workflow in which agents perform, coordinate, and support the workflow [Huhns/Singh 98]. In an agent-based workflow system, there exist different agent types that manage the workflow. A process instance agent is responsible for controlling one process instance. Newer architectures further split the functionality of a workflow system: task agents, which are mobile and migrate to subjects which perform some tasks of the workflow [Hawryszkiewycz/Debenham 98], and personal agents which act as an interface between the subject and other agents are introduced.

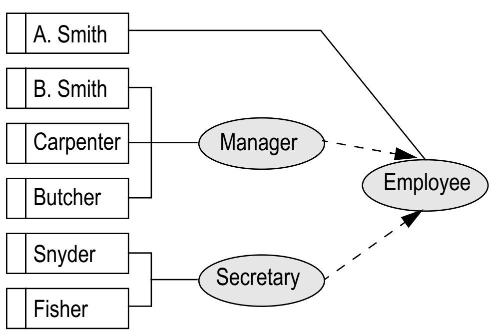


Fig. 1: Sample role definition and hierarchy

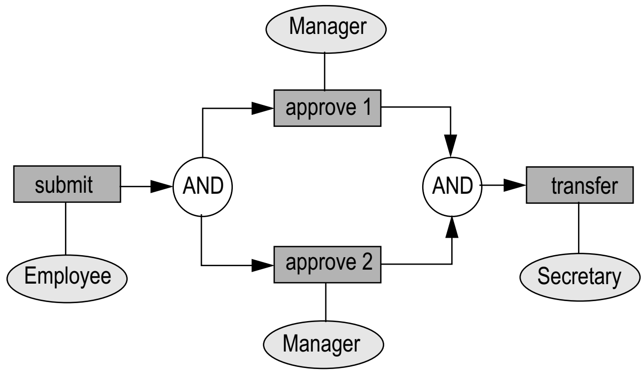


Fig. 2: Example of a process definition

3 Example Workflow Environment

This section gives an example which will be used throughout the paper for illustration purposes.

Six persons, also referred to as subjects, *A. Smith*, *B. Smith*, who are brothers, *Carpenter*, *Butcher*, *Snyder*, and *Fisher* are working in a company. Figure 1 shows these subjects together with their assigned roles. Note that every *Manager* (or *Secretary*) is an *Employee*, too. E.g. *Butcher* is able to activate the manager or the employee role. The partial order of roles builds up a so called *role hierarchy* – a well known modelling approach [Scheer 94].

Figure 2 shows the process definition for a travel expense claim. The process starts when an employee submits a travel expense claim. Two Managers must approve this claim before the money transfer is done by a secretary. Note that the tasks are partially ordered, e.g. *submit* precedes all other tasks but no order is possible between *approve 1* and *approve 2*.

4 Architecture of the Agent-based Workflow System

The architecture of the proposed agent-based workflow system consists out of the following four agent types:

Process Instance Agent (PIA) The PIA is created by a subject, which has to provide a complete and correct process definition. The process instance agent represents and manages an instance of a workflow (according to the process definition given) and controls its whole execution. In the example a PIA is created for each travel expense claim of an employee (e.g. Claim 157 of Butcher).

Task Instance Agent (TIA) The TIA is responsible for one task in a process instance. It is created by the PIA and has to search for a subject, deliver the task description and objects items to the subject and the results back to the PIA. Referring to the example, a TIA is created for the transfer task of Claim 157 of Butcher.

Worklist Agent (WLA) The WLA stores a mapping of all subjects and their assigned roles (cf. Figure 1).

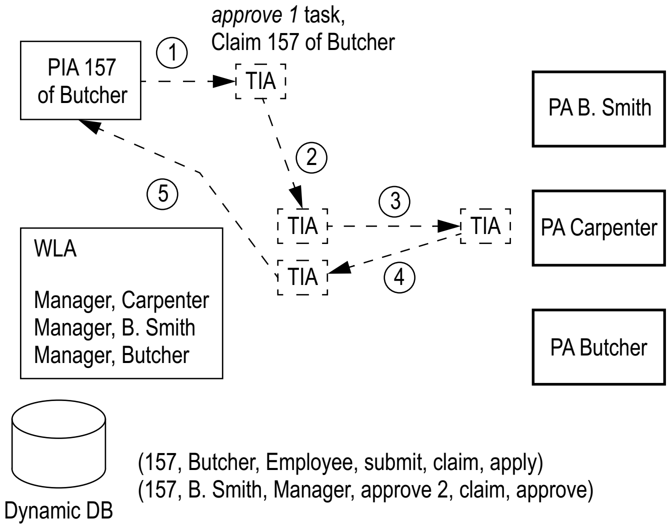


Fig. 3: Life cycle of a TIA

Personal Agent (PA) A personal agent is the interface between the subject and the incoming TIAs. It is immobile, controls the incoming task requests from the TIAs, and coordinates the communication between subject and TIA.

The basic idea in the handling of a workflow task is to create a TIA for each task in the underlying process instance. The PIA creates all TIAs according to the temporal and logical prerequisites of the workflow. E.g., the parallel execution of several tasks can be achieved by the creation of several TIAs simultaneously.

When the PIA creates a TIA for fulfilling task *t*, all and just those objects and privileges which are needed for the task's execution are instantiated as part of the TIA. Next, the TIA has to find a subject for task *t*. Therefore, it migrates to the WLA to get a list of possible subjects to interact with *t*, qualifying by means of role allocations and possibly other constraints. The TIA chooses a subject *s* randomly from the list and migrates to *s*: This is another security feature because the choice of the TIA is not predictable and therefore fraud is complicated. Clark and Wilson [Clark/Wilson 87] identified SoD as one of the two major mechanisms that can be implemented to ensure data integrity. They purposed the random selection of task participants in order to ensure that any attempted conspiracy is inherently unsafe. More elaborate choices are possible but are not further investigated as part of this paper.

After the migration of the TIA to the selected subject, the PA of *s* is informed that a new task is waiting to be executed. The PA must now inform the subject, for example by showing a message on the screen or playing a sound sample. Then, the subject performs the task using the objects and privileges which will be provided by the TIA. When the task is done, the TIA migrates back to the WLA to pass information to the dynamic database of WLA. This information will contain *s*, *t*, the objects and privileges used plus other information like time stamps. Finally, the TIA migrates back to the PIA to pass control flow related information.

Fig. 3 illustrates the above procedure by means of the travel expense claim example. The PIA instantiates the task *approve1* in the process instance 157 of Butcher (1). The TIA migrates to the WLA to get a list of potential subjects who qualify for the execution of the task (2). The WLA creates the subject list based on rules and on the dynamic database (cf. Section 5.3). In this example the list contains only one subject, Carpenter, because the dynamic database shows that B. Smith has done the other approve task and Butcher cannot approve his own claim. Therefore, the TIA migrates to PA of Carpenter and the task is executed. Next, the TIA migrates back to the WLA to pass information to the dynamic database. Finally, the TIA migrates back to the PIA.

5 Security Aspects

The focus of the paper is to show that SoD and RBAC mechanisms are feasible for implementation in an agent-based workflow system. Therefore, this section gives a brief introduction to RBAC and SoD. Then, a formal model is proposed facilitating the implementation of access control (by means of RBAC) and integrity (by means of SoD) in an agent-based workflow environment.

5.1 RBAC

In a workflow environment, usually tasks are not linked directly to subjects. The concept of *roles* forms a middle layer between the subjects and the tasks, cf. Figure 4. As an example consider the *Manager* role in the travel claim example. Furthermore, access rights are enforced on roles and not on subjects. This simplifies the security administration. There is a considerable amount of research about RBAC going on.[1]

5.2 SoD

"SoD is a policy to ensure that failures of omission or commission within an organization are caused only by collusion among individuals and, therefore, are riskier and less likely, and that chances of collusion are minimized by assigning individuals of different skills or divergent interests to separate tasks" [Gligor et al. 98]. In a workflow context, SoD has to be divided and extended into static SoD (SSoD) and dynamic SoD (DSoD). SSoD enforces certain rules during build time of the workflow, i.e. before process instances of the workflow are instantiated. Example: The process definition in Figure 2 requires that different tasks are performed by different roles (the *transfer* task by a secretary and the *approve 1* task by a manager). SSoD rules are applied to a process definition to guarantee their enforcement. In contrast DSoD can only be enforced during run time of the workflow, i.e. during the execution of a single instances of a process. DSoD can be enforced on different layers. Consider the following examples from the scenario introduced in Section 3:

1. A manager should not be allowed to submit for (first task) and approve (in a later task) his own travel expense claim. In this case DSoD should be enforced on the role and subject layer.

1. http://www.acm.org/sigsac/rbac2000.html

to Contents, page 1

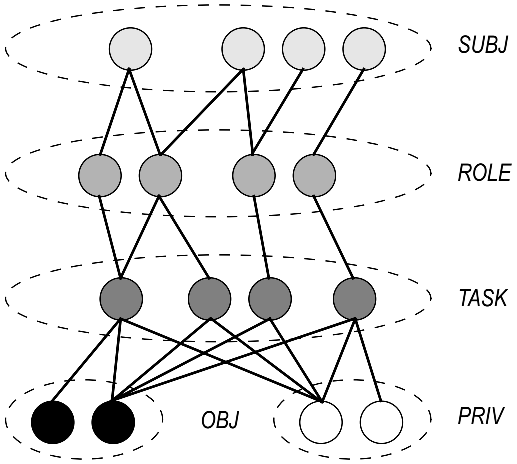


Fig. 4: Sets and mappings in the DSoD model

2. The following example is on a subject layer: Two brothers (A. Smith and B. Smith) are working in the same company. To prevent fraud, one of the brothers should not be allowed to approve the travel expense claim of his brother.
3. As a third example considers a document which is used in a first task by a subject in a certain role. In the travel claim example there could be the policy that the subject should not get any privileges to this document for the remaining tasks in the process instance even in a different role. This could be necessary to prevent any (ex post) manipulation of the document. Suppose Snyder submits a travel expense claim in her employee role. Then, Snyder can not transfer the money even if her secretary role is activated. This example illustrates SoD on an object, subject and role layer.

There are more complex SoD policies [Ahn/Sandhu 99, Bertino et al. 99] which fall out of the scope of this paper. The model which will be introduced in the following subsection allows for defining dynamic policies on role, subject, task, object and privilege layer. To illustrate the model, the three examples given above will be used.

5.3 A model for DSoD

Our model uses set theory and functions using the following five sets:

SUBJ The set of all persons who are capable of executing a task.

ROLE The set of all roles in a process definition.

TASK This set includes all tasks that are defined in the process definition.

OBJ The set of all objects (e.g. a text document) which are needed to execute the tasks.

PRIV The set of privileges. A privilege is used on an object, for example a read permission on a text document.

In the workflow definition it has to be defined which subjects are allowed to activate which roles. Therefore, the function *R* maps *SUBJ* to the power set of *ROLE*.

$$R{:}SUBJ \rightarrow 2^{ROLE}$$

In a next step, all task definitions have to include the role that is allowed to execute the task. The following function is defined:

$$T{:}TASK \rightarrow 2^{ROLE}$$

Finally, a function is needed to associate tasks with objects and privileges:

$$P{:}TASK \rightarrow 2^{OBJ} \times 2^{PRIV}$$

Figure 4 shows the relationship between the five sets and the three functions. Information constrains in the above mentioned sets must be defined in the process definition. At run time, it is possible to implement DSoD using these sets together with a natural number, indicating the process instance.

Consider a tuple consisting of the Cartesian product of the following six sets where N is the set of all natural numbers and n indicates the process number:

$$(n, s, r, t, o, p) \in N \times SUBJ \times ROLE \times TASK \times OBJ \times PRIV$$

Not all of these tuples are meaningful. Therefore, the notion of soundness is introduced. A tuple (n, r, s, t, o, p) is *sound* if and only if

1. $r \in R(s)$,
2. $r \in T(t)$, and
3. $(o, p) \in P(t)$

hold. The first inclusion says that the subjects s should be allowed to activate role r. Next, the role r should be allowed to execute task t. Finally, the object o and the privilege p should match with the task t. Note that the tasks have to be in a chronological and logical order. We therefore assume that (TASK, ≤) is a partially ordered set. $t_1 \leq t_2$ says that t_1 is executed in parallel with or before t_2.

The WLA uses a database called DB_{dyn} for storing all tuples and rules to enforce DSoD. The rules for our examples above are:

1. A manager who is not allowed to approve his own claim. n is a process instance id, s a subject, r the manager role, t_1 is the submit task, t_2 the approve 1 or 2 task, o is the travel claim and p are the submit privilege.
$$(n, s, r, t_1, o, p) \in DB_{dyn} \Rightarrow (n, s, r, t_2, \#, \#) \notin DB_{dyn}$$
\# symbolizes any possible element of the corresponding set.

2. Consider the "brother" example, where s_1 is A. Smith and s_2 is B. Smith, t_1 is submit, t_2 is approve 1 or 2 task, o is the travel claim document, p_1 is the submit and p_2 the approve privilege.
$$(n, s_1, \#, t_1, o, p_1) \in DB_{dyn} \Rightarrow (n, s_2, \#, t_2, o, p_2) \notin DB_{dyn}$$

3. Last example: s is Snyder, r is Employee, t_1 is submit and o is the travel claim.
$$(n, s, r, t_1, o, \#) \in DB_{dyn} \Rightarrow (n, s, \#, t_2, o, \#) \notin DB_{dyn}$$
for all $t_1 < t_2$

It is very important to check the rules for consistency. Contradicting rules may ruin the whole SoD mechanism. Also it might be possible to combine several rules to a single rule.

This process is called *pruning*. For more details see [Bertino et al. 99].

6 Implementation of DSoD in an Agent-based Workflow System

Section 4 introduced an architecture for an agent-based workflow system, Section 5.3 a formal model for DSoD rules. This section shows how these parts are interrelated. For illustration refer to Figure 3 to see which information from the DSoD model is used when and by which agent types.

The PIA has control over the whole process execution, i.e. the partially ordered set of tasks *TASK*, the roles *ROLE*, objects *OBJ*, and privileges *PRIV* associated with each task. In step (1) the PIA instantiates a TIA for task t and passes information about the task. This information consists of $P(t)$ (the privileges, objects) and $T(t)$ (the roles) associated with t. Note that only the privileges and objects are passed which will be needed by the subject to perform the task implementing the security principle of least privilege.

The TIA migrates to the WLA and passes $T(t)$ to the WLA. The list of possible subjects is created by checking the elements of $T(t)$ and $R(s)$ for all tasks and subjects.

In step (3) the TIA decides randomly which subject performs the task and gives this information back to the WLA. The WLA will generate a subject list $[s_1,..., s_n]$ after enforcing the DSoD rules based on the entries in the dynamic database DB_{dyn}. This list consists of all subjects capable of executing the task and will usually contain more than one subject. The TIA uses a random function rand which maps the input n to a random number in the set $\{1, \ldots,n\}$. The subject chosen will be $s_{rand(n)}$. This practice further decreases the possibility of fraudulent activities since the outcome of the assignment is not predictable.

Now the WLA blocks the subject for this instance. This is necessary for parallel execution, otherwise the DB_{dyn} checking could fail. Before the TIA migrates to the subject the WLA inserts new tuples in the dynamic database. A tuple is (n, r, s, t, o, p), where n is the process instance number, r the role associated with task t, s the subject which performs t, and p the privilege used on object o. Several tuples can be inserted into DB_{dyn} for t, e.g. if several objects are used in the task. When the TIA arrives at the subject, the PA of s gets the objects plus privileges contained in $P(t)$. Depending on the task, one or more objects may be created which have to be transferred via the TIA to the PIA. (4) takes the TIA back to the WLA where the subject is unblocked.

As a final step (5) the TIA moves back to the PIA where the objects are passed over. Finally, the TIA is deleted by the PIA. If a complete process instance has finished (in the example, after the transfer task), the PIA can initiate a garbage collection at the WLA which removes all tuples from this process instance from DB_{dyn}. This information together with the objects created during the execution can be stored in an archive. Finally, the PIA is deleted, too.

7 Conclusion

This paper introduced an agent-based workflow environment consisting of four different agent types. The following table shows the four agent types and their most important characteristics.

	Mobility	Intelligence	Security	Instance
PIA	-	+	+	+
TIA	+	+	+	++
WLA	-	-	-	-
PA	-	-	-	-

The mobility of a TIA is rated high (+), since a TIA has to do several migrations during its life. The TIA will be an “intelligent” agent because several decisions such as the choosing of a subject from the subject list has to be done. The TIA furthermore enforces the strict least privilege principle (Security +). A TIA has to be created for every task in a process instance (marked ++ in the Instance column). The other agents are interpreted similarly.

This workflow environment was used to enforce different security features:

1. (Dynamic) access control: The access to objects is restricted. This is done dynamically because TIAs are created based on the state of the process instances.
2. Strict least privilege: A subject will just receive privileges to objects which are needed for the execution of a task.
3. SoD: The major focus of the proposed architecture is SoD. Its dynamic variant DSoD is realized through the interaction of PIA and WLA.
4. Random choice of subjects: This practice further decreases the possibility of fraudulent actions.

The work suggests that agents are a valuable resource in implementing security features in workflow environments.

Future work will deal with the development of a prototype to validate the presented architecture. Furthermore, the SoD model will be extended, e.g. to enforce rules between different workflow instances, e.g. when confidential data could be collected in different workflow instances by the same subject.

References

[Ahn/Sandhu 99]
Gail-Joon Ahn and Ravi Sandhu. The RSL99 Language for Role-Based Separation of Duty Constraints. In Proceedings of 99 ACM Conference on Role Based Access Control, 1999.

[Bertino et al. 99]
Elisa Bertino, Elena Ferrari, and Vijayalakshmi Atluri. The Specification and Enforcement of Authorization Constraints in Workflow Management Systems. ACM Transactions on Information and System Security, 2(1):65–104, February 1999.

[Crystaliz et al. 97]
Crystaliz, Inc., General Magic, Inc., GMD Focus, and IBM Coorp. Mobile Agent Facility Specification. Technical report, OMG, 1997.

[Clark/Wilson 87]
David D. Clard and David R. Wilson. A Comparison of Commercial and Military Computer Security Policies. IEEE Symposium on Security and Privacy, pages 184–194, 1987.

[Ferber 99]
Jacques Ferber. Multi-Agent Systems: An Introduction to Artificial Intelligence. Addison Wesley Publishing Company, 1999.

[Gligor et al. 98]
Virgil D. Gligor, Serban I. Gavilla, and David Ferraiolo. On the

to Contents, page 1

formal definition of separation-of-duty policies and their composition. IEEE Computer Society, IX:172–183, 1998.

[Georgakopoulos et al. 95]
Dimitrios Georgakopoulos, Mark Hornick, and Amith Sheth. An Overview of Workflow Management: From Process Modelling to Workflow Automation Infrastructure. Distributed and Parallel Databases, 3:119–153, 1995.

[Hawryszkiewycz/Debenham 98]
Igor Hawryszkiewycz and John Debenham. A Workflow System Based on Agents. In Gerald Quirchmayr, Erich Schweighofer, and Tervor J. M. Bench-Capon, editors, Database and expert systems applications, volume 17, page 688ff. Springer, 1998.

[Huhns/Singh 98]
Michael N. Huhns and Munindar P. Singh. Workflow Agents. IEEE Internet Computing, July/August 1998.

[Leymann/Altenhuber 94]
F. Leymann and W. Altenhuber. Managing Business Processes as an Information Resource. IBM Systems Journal, 33(2):326–348, 1994.

[Lange/Oshima 99]
Danny B. Lange and Mitsuru Oshima. Seven Good Reasons for Mobile Agents. Communications of the ACM, 42(3):88, March 1999.

[Scheer 94]
August-Wilhelm Scheer. Business Process Engineering. Reference Models for Industrial Enterprises. Springer Verlag, Berlin, 1994.

Secure Server-Side Java Applications

Ulrich Ultes-Nitsche

Java Servlets offer a convenient way to implement Java programs that run on the server side. Being exposed to public access via the Internet, they are equipped with security mechanisms to protect the Internet server from various types of threats. I will discuss in this article the security features provided by Java Servlets. In addition, I will describe briefly the additional mechanisms that need to be put into place to create secure server-side Internet applications.

Keywords: secure Internet applications, Java servlets, secure sockets layer protocol (SSL)

1 Introduction

The past years have seen the rise of server-side Java applications, known as Java Servlets. Servlets are used to add increased functionality to Java-enabled servers in the form of small, pluggable extensions. When used in extending web servers, Servlets provide a powerful and efficient replacement for CGI (Common Gateway Interface) and offer many significant advantages [Sun 98]. These advantages include:

- **Portability**: Java Servlets are protocol and platform independent and as such are highly portable across platforms and between servers. The Servlets must conform to the well defined Java Servlet API which is already widely supported by many web servers.
- **Performance**: Java Servlets have a more efficient life cycle than either CGI or FastCGI scripts. Unlike CGI scripts, Servlets do not create a new process for each incoming request. Instead, Servlets are handled as separate threads within the server. At initialisation, a single object instance of the Servlet is created that is generally persistent and resides in the servers memory. This persistence reduces the object creation overhead. There are significant performance improvements over CGI scripts in that there is no need to spawn a new process or invoke an interpreter [Hunter 98]. The number of users able to use the system is also increased because fewer server resources are used for each user request.
- **Security**: As the security of server-side Internet applications will be the major concern of this article, it will be discussed in a separate section subsequently. However, it should be mentioned here that, by being a derivate of the full Java programming language, Java Servlets inherit all Java security mechanisms plus some Servlet specific features.

As Servlets aim at being used where CGI scripts were used historically, I will compare mainly the two paradigms. As not only will the server-side application have to meet reasonable security requirements, but the entire Internet-based solution as a whole, I will in the second part of this article discuss briefly the use of SSL to secure communication with the application.

2 Security of Java Servlets

In this section, I discuss the security features of Java Servlets. The Java language and Java Servlets have improved security over traditional CGI scripts both at the language level and at the architecture level:

Language Safety

As a language Java is type safe and handles all data types in their native format. With CGI scripts most values are treated and handled as strings which can leave the system vulnerable. For example, by putting certain character sequences in a string and passing it to a Perl script, the interpreter can be tricked into executing arbitrary and malicious commands on the server.

Java has built-in bounds checking on data types such as arrays and strings. This prevents potential hackers from crashing the program, or even the server, by overfilling buffers, in case of not properly written code. For example, this can occur with CGI scripts written in C where user input is written into a character buffer of a predetermined size. If the number of input characters is larger than the size of the buffer, it causes a buffer overflow and the program will crash. This is commonly known as stack smashing.

Java has also eliminated pointers and has an automatic garbage collection mechanism which reduces the problems associated with memory leaks and floating pointers. The absence of pointers removes the threat of attacks on the system where accesses and modifications are made to areas of server memory not belonging to the service process.

Dr. Ulrich Ultes-Nitsche is a lecturer at the University of Southampton, U.K. He studied Computer Science and Physics at the University of Frankfurt. Before joining Southampton, he was a junior scientist at GMD Darmstadt, a visiting researcher at the University of Liège, and a research assistant at the University of Zürich. In 1995 he received a DAAD HSP II/AUFE fellowship and in 1997 he received the Best Student Paper award of the 16th Annual ACM Symposium on Principles of Distributed Computing (PODC'97), Santa Barbara, USA. His research interests comprise security, formal methods, and automated verification. His web-page is at http://www.ecs.soton.ac.uk/ ~ uun.

to Contents, page 1

Finally, Java has a sophisticated exception handling mechanism, so unexpected data values will not cause the program to misbehave and crash the server. Instead an exception is generated which is handled by the program that usually terminates neatly with a run time error [Garfinkel/Spafford 97].

Security Architecture

Java Servlets have been designed with Internet security issues in mind and mechanisms for controlling the environment in which the Servlet will run have been provided.

CGI scripts generally have fairly free access to the server's resources and badly written scripts can be a security risk. CGI scripts can compromise the security of a server by either leaking information about the host system that can be used in an attack, or by executing commands using untrusted or unchecked user arguments. Java significantly reduces these problems by providing a mechanism to restrict and monitor Servlet activity. This is known as the Servlet Sandbox. The Servlet Sandbox provides a controlled environment in which the Servlet can run and uses a security manager to monitor Servlet activity and prevent unauthorised operations. There are four modes of operation that include trusted Servlets, where the Servlet has full access to the server resources, and untrusted Servlets which have limited access to the system.

JDK 1.2 contains an extension to its security manager, the access controller. The idea behind the access controller is to allow more fine-grained control over the resources a Servlet can access. For example, instead of allowing a Servlet to have write permission to all files in the system, write permission can be granted for only the files required by the Servlet for execution [Hunter 98].

However, Java-based servers are still vulnerable to denial of service attacks where the system is bombarded with requests in order to overload the server resources. This approach invokes so many Servlet instances that all the server resources are allocated. This can impact all the services supported by the server. However, the effects of this can be reduced by specifying an upper limit on the number of threads that can be run concurrently on the server. If all the threads are allocated, that particular service can no longer be accessed, but because the server still has resources left to allocate, the rest of the services are still available.

3 Secure Server Access: SSL

In this section I discuss the security features of the network connection to the Servlet. I do so by focusing on the use of SSL. The secure sockets layer protocol (SSL) is designed to establish transport layer security with respect to the TCP/IP protocol stack. Version 3 was published as an Internet draft document [Freier et al. 96] by the IETF (Internet Engineering Task Force). The use of SSL will be introduced briefly along the lines of [Stallings 98].

3.1 The Protocol Stack

The transport layer part of SSL, the SSL record protocol, sits on top of TCP in the Internet protocol stack. It is accessed by an upper layer consisting of the hypertext transfer protocol (HTTP) and different parts contributing to SSL: SSL handshake protocol, SSL change cipher spec protocol, and the SSL alert protocol, used to set up, negotiate, and change particular security settings used by the SSL record protocol. Schematically, the SSL architecture is presented in Figure 1.

Besides the record protocol, all other protocols are needed to establish a secure connection to the server-side application. They perform authentication and access-control steps, mainly without user involvement. After a connection is established, the use of SSL (via the record protocol) is completely transparent to the user: she/he still accesses only http.

Security Features of SSL

SSL allows for different security features being chosen. First of all, different encryption algorithms can be used to produce ciphertexts and authentication messages. For authentication, different hash algorithms can be negotiated. SSL can also use X509.v3 peer certification [Garfinkel/Spafford 97]. Using a session identifier, active states of SSL are identified, where a state consists of a number of keys involved in the session, both on the server and on the client side, and sequence numbers to count the messages exchanged. By using these different parameters, SSL sets up a session configuration that then allows for ensuring integrity, confidentiality, and authentication depending on the set up parameters.

Benefits of Using SSL

Other authentication mechanisms for Internet applications are HTTP authentication [Franks et al. 97] and digest authentication [Berners-Lee et al. 96]. They only provide a basic authentication mechanism, lacking mechanisms for confidentiality and integrity. In contrast, SSL offers the full range of security mechanisms needed to establish a secure session. It is particularly the secure session concept that makes SSL appealing to be used in combination with Java Servlet implementations. The secure session last as long as a connection to the server-side application exists. Communication with the user is then based on html-documents sent to her/him using http. As soon as an SSL session is established, it is completely transparent, preserving http as a stateless protocol.

HTTP authentication or digest authentication would establish a suitable mechanism for secure server-side applications only if TCP/IP were secured, either by equipping IPv4 with IPsec or by running IPv6 (IP version 6 is the latest IP version and includes IPsec) [Garfinkel/Spafford 97]. IPsec is a low layer protocol in the Internet protocol stack for end-to-end

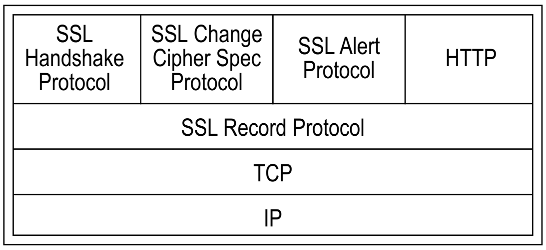


Fig. 1: SSL within the Internet protocol stack [Stallings 98].

confidentiality of the transmission of IP-packages. However, SSL is a security protocol widely deployed such that its choice appears to be more practical than choosing any combination of IPsec/IPv6 with HTTP authentication/digest authentication.

4 Conclusions

To summarize, Java Servlets offer an appealing way to implement secure server-side application when combined with SSL. The run-time environment of Java Servlets offers reasonable built-in security features. The use of Servlets is described in the framework of a web-hosted e-mail system in [Hepworth/Ultes-Nitsche 99] and in connection with a context-dependent access control to medical data in [Ultes-Nitsche/ Teufel 00].

If one had to build a secure Internet application in which the use of SSL appears insufficient to provide the security requirements, a combination of Java Servlets on the server side and Java Applets on the client side can be useful. In such a scenario, the server-side application would not communicate a browser by sending html-documents but by communicating with an Applet. This would allow the use of proprietary security protocols that may enhance the security of the system. For instance, mouse-movement can then be used to create nonces (randomly created protocol parameters). However, using Java Servlets to implement server-side applications and SSL in the client/server communication, as discussed in this article, will offer reasonable security mechanisms sufficient for most Internet applications.

Acknowledgements

I would like to acknowledge gratefully the use of material from joint work with Eleanor Hepworth and Stephanie Teufel. I am very thankful to both of them.

References

[Berners-Lee et al. 96]
T. Berners-Lee, R. Fielding, and H. Frystyk. Hypertext Transfer Protocol – HTTP/1.0, Internet Standards Request For Comments number 1945, May 1996

[Franks et al. 97]
J. Franks, P. Hallam-Baker, J. Hostetler, P. Leach, A. Luotonen, E. Sink, and L. Stewart. An Extension to HTTP: Digest Access Authentication, Internet Standards Request For Comments number 2069, January 1997

[Freier et al. 96]
A. O. Freier, P. Karlton, and P. C. Kocher. The SSL Protocol Version 3.0. Internet Draft. Netscape, Transport Layer Security Working Group, November 1996. See http://home.netscape.com/eng/ssl3/draft302.txt.

[Garfinkel/Spafford 97]
S. Garfinkel and G. Spafford. Web Security and Commerce, O'Reilly and Associates 1997.

[Hepworth/Ultes-Nitsche 99]
E. Hepworth and U. Ultes-Nitsche. Security aspects of a Java-Servlet-based web-hosted e-mail system. In: J.H.P. Eloff, L. Labuschagne, R. von Solms, and J. Verschuren (eds.). Proceedings of the 7th IFIP Working Conference on Information Security Management & Small Systems Security, Amsterdam, Kluwer Academic Press, 1999.

[Hunter 98]
J. Hunter. Java Servlet Programming, O'Reilly and Associates, 1998.

[Stallings 98]
W. Stallings. Cryptography and Network Security. Prentice Hall, New York, second edition, 1998.

[Sun 98]
Sun Microsystems. Java Servlet API Whitepaper, 1998.

[Ultes-Nitsche/Teufel 00]
U. Ultes-Nitsche and S. Teufel. Secure access to medical data over the Internet. In: H. R. Hansen, M. Bichler, and H. Mahrer (eds.). Proceedings of the 8th IFIP European Conference on Information Systems, Vienna University of Economics and Business Administration, 2000.

to Contents, page 1

Quality of Service in web applications

Antoni Drudis

The performance of a large distributed system depends not only on the performance of its components but also on the way that these components interact with each other. In the case of the web, where the designer has little or no control over most of the components of the solution, applications have to provide tools to manage the performance of the solution. Quality of service was initially formalized for the network traffic and it is being extended to the application level. This article introduces the main concepts of web server performance characterization and describes a sockets-based implementation on several operating systems.

Software design is an evolving discipline where every major milestone is the consolidation of existing technologies and the solutions for new application requirements in terms of resilience, development costs, and functionality. Software modularization led to client-server architectures by delegating application intelligence to distributed objects. In turn, client-server architecture provides the foundation for web applications to address the requirements on distributed transactions over the Internet.

Since functions centralized in the mainframe were delegated to autonomous servers in the client-server paradigm, new services were built. For example, the need for a uniform resource identification policy for objects led to name servers, and the need to protect services from unauthorized access led to secure services and encryption techniques. From the user's perspective, client-server architectures provided a lower-cost, flexible, and scalable solution where services and processing power could be added or modified without a major rewrite of the application software.

The distributed nature of web applications requires a flexible, scalable, and fault-tolerant architecture that is not adequately served by traditional transaction-processing engines. While the basic building blocks of software tools have not changed much over time, the way these blocks are put together reflects the new requirements on the overall solution. The initial applications on the web can be conceptually seen as a client-server model where web servers deliver data to browsers and other clients such as applets and search robots. What is new with the web is the scale of the global network and the patterns of server access. Concurrent users are not counted in thousands anymore, but in tens or hundreds of thousands, the environment is heterogeneous and constantly changing, transactions are stateless and difficult to track, and the process may integrate applications from different environments.

Scaling can be achieved by upgrading the processing platforms to faster systems but more so by distributing services and replicating servers. That strategy not only provides increasing computing power without bringing down the whole network but it also improves the resilience of the solution and allows for planned and unplanned down time.

The standardization of application components facilitates the interoperability of servers and clients on the web. The transport protocol is a de-facto standard, markup languages are being standardized, and file formats such as PDF or programming languages such as Java are being used across operating systems and computing platforms.

The classic paradigm of transaction processing is being extended to the web. On servers and mainframes, transaction monitors usually are based on the assumption that the client or the remote terminal will send well-behaved transactions. These transactions are short, serializable, most frequently stateful, and homogeneous. In the initial instances of the web usage to publish data, that concept fits the requirements of the web server: users send requests to get data from a file or out of a software module, and the server processes each transaction on a best-effort basis. On transaction processing engines, best-effort optimization is reached by minimizing the waiting time for each transaction.

Instead, web servers for database access, such in business to business or business to consumer transactions, introduce the concept of *session* where a set of transactions has to be processed under the same assumptions about the user environment. In a session clients may pause for long periods of time while the user interacts with the application.

For example, a user accesses a book catalogue on a web bookstore, creating a transaction. The web server communicates with the application server to request database entries that satisfy the query conditions. Then, the user selects one or more books to buy them. While the user is waiting, the store could have changed the price of the book, but the buyer still expects

Antoni Drudis is a software architect at the Telecommunications Infrastructure Division of Hewlett-Packard Company in Cupertino, California. He worked on the HP-UX and OSF operating systems and was a member of the X/Open and International Standards Organization. Antoni designed and developed the management server for WebQoS. He holds computer science and industrial engineering undergraduate degrees and a PhD in engineering by the Universitat Politecnica de Catalunya in Barcelona.

the old price to be applied to this sale. At this point, the buyer places the order and the web server communicates with the application server which, in turn, may place an order to the publisher's server, request space in the warehouse, start a shipping order, verify and charge the customer's credit card, and send an e-mail confirmation. All these individual transactions on different servers and subnets form a single session because all of them have to use a uniform set of data from the environment. While there are many transaction engines commercially available, sessions are managed by customized applications.

The need for immediate solutions for the web growth requirements has led to solutions in the areas of state preservation, scaling processing power, and application resilience. The Java language is making its inroads as a pervasive computing environment at the client side, thus reducing the need for bandwidth and server processing. Efficient queuing management tied to a multithreading paradigm provides a first level of smoothing out the frequency of variations in client accesses, and solutions at the client side such as cookies and the server side such as application servers provide some state preservation in a fundamentally stateless environment. Some specialized applications such as publishing – where data is stable, access is repetitive, and the cost of local storage is smaller than cost of bandwidth – can use data and server caching and platform replication as a means to reduce the current strain on servers.

Replication and caching are further utilized with static data to increase the tolerance of the overall network to both planned and unplanned maintenance tasks. Communication from the client is no longer an all or nothing proposition but a guaranteed operation when the access parameters stay in a given range. System and network administrators can configure the resources to provide a fail-over mechanism that prevents catastrophic consequences when a key component fails.

While optimizing individual components of the e-commerce solution may prove to be a good short-term strategy for product vendors to accelerate their market penetration, solution integrators are still faced with the formidable task of translating these component-level performance improvements into a predictable end-user experience.

Since network communications are bound by the slower component, solution architects often design redundant solutions with replicated components such as network segments and computer systems. These topologies scale only if the appropriate algorithms run on the nodes to divert the traffic into the most efficient branch and quality of service needs to be built in the protocol so the appropriate priority is assigned to data packets travelling the network.

The web introduces a new challenge to application designers. Users are no longer trusted and well-trained employees but anonymous clients who may engage in non-cooperative or destructive behaviour that increase resource consumption and endanger the integrity of the database. Once again, the distributed and stateless nature of the application makes it difficult to maintain efficient and secure communications between the server and its clients.

Commercial users of web-based services cannot longer accept ambiguous quality of service promises. Instead, they require service level agreements with the solution providers to guarantee a predictable web experience for the users who access the web site to perform financial transactions.

A service level agreement (SLA) describes the service in terms of number of concurrent users, types of resources available to these users, maximum bandwidth, response time, down time for maintenance or upgrades, storage capacity, backup policies, security tools and processes, accounting, and other services. What differentiates an SLA from other types of contracts is its focus on results instead of focusing on the resources spent to provide these results. For example, SLAs have penalty clauses associated to the lack of performance on any of the guarantees specified in the contract.

Service level agreements between the solution providers and the sponsors of the e-commerce site are translated into service level objectives (SLO) that describe the desired behaviour for web servers, application servers, and back-end software.

The SLA focuses on the user experience by measuring variables such as the time to download a web page or the number of users who may simultaneously access a web server without impacting the response time more than a given percentage. The SLO may specify parameters such as download time per megabyte, the response time to an HTTP request, or the number of user-generated transactions that the web server can forward to an attached application or database server. Both SLAs and SLOs facilitate the business specialization of Internet companies. Content providers produce the actual data the user is interested in. Application service providers (ASP) and web hosting companies allow content providers to concentrate on their speciality by guaranteeing service to customers, and Internet service providers (ISP) allow ASPs and clients to access the network with guaranteed bandwidth.

For the application developer, the goals for quality of service affect the way the application processes its transactions. Instead of using a best-effort method that avoids the analysis of the transaction priority to reduce process time, transactions are queued and scheduled according to the request priority.

Some commercial application servers and web server platforms provide support for quality of service capabilities. For example, Hewlett-Packard Web Quality of Service product family (WebQoS) is a set of software tools that address the configuration and management of service level objectives for web servers running on the platform. Using WebQoS, administrators define the expected operational parameters for the e-commerce site as well as the corrective actions to be executed. When these objectives are not met, the WebQoS tools automatically adjust the behaviour of the servers and inform the user about the status and operational parameters of the domain.

The process to implement WebQoS is simple. First, the analyst observes the access patterns to the web server. For example, what percentage of static data is accessed which pages consume more resources, and how much time users spend on each page. The set of measurements leads to a baseline that reflects the time of day, day of week, and seasonal differences in access patterns. The analysis is done with the help of logging files, network management tools such as OpenView or Fire-

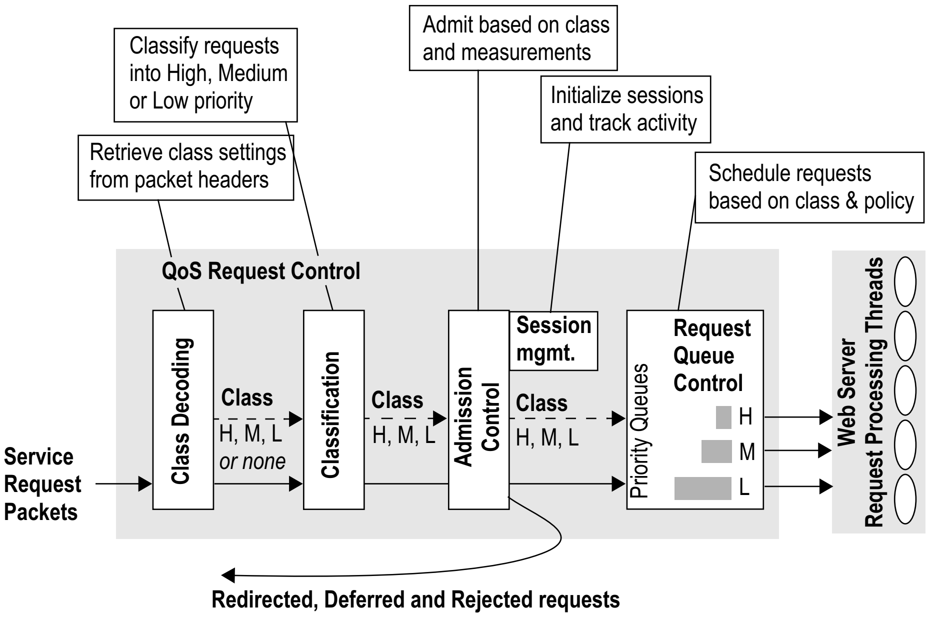


Figure 1: Classification provides the mechanism by which relative priorities to user requests can be assigned

hunter, and field measurements conducted by third parties such as Resonate.

Next, the analyst measures the impact of changes in the environment and predicts the requirements on the web server. For example, when company web site is shown on television or the press, there is a surge of users who want to check the site by themselves. The same is true when there is an aircraft accident, natural disaster or other unexpected events.

Then, the analyst investigates the effect of internal incidents such as network bottlenecks, proxy malfunctions, crashes of servers in the server farm or hardware and software upgrades. The result of that analysis is an estimate of the parameters that characterize quality of service from the customer's point of view, noting that the behaviour of the user often compounds to the busy server problems. When users find the response time unacceptable, a typical reaction is to press the reload icon. The effect on the server is to have yet one more transaction to process and the need to discard all HTTP requests is no longer valid.

From the web server's point of view, when the number of users exceed the maximum capacity threshold, the resources of the system are not spent in processing transactions but dealing with the unexpected load. Response time grows exponentially and the server becomes inoperative.

The analyst formalizes at this time the service level objectives of the server. For example, an SLO may specify the response time of each HTTP request, measured from the time the server reads an HTTP request up to the time the server sends the answer to that request back to the client. Other SLOs specify the maximum number of concurrent sessions, the average CPU load of the system where the web server is running on and queue length, which is the number of requests that have not yet been processed by the server.

Since SLOs are defined in terms of thresholds, the administrator has to specify what to do when the threshold has been exceeded. One or more corrective actions may be assigned to each SLO to define the behaviour of the web server. For example, the request could be rejected, deferred, or redirected to another URL in the same or another server.

Once the operational parameters of the web server have been defined, the analyst works with the content provider to determine the priority of web accesses. For example, while a given merchant may provide better response time to customers who are actually buying books on the web, another merchant may provide better response time to users who browse through the catalogue. Similarly, while a merchant may want to maximize the number of simultaneous users who access the web site, another might want to put a cap to guarantee that all orders can be delivered as promised.

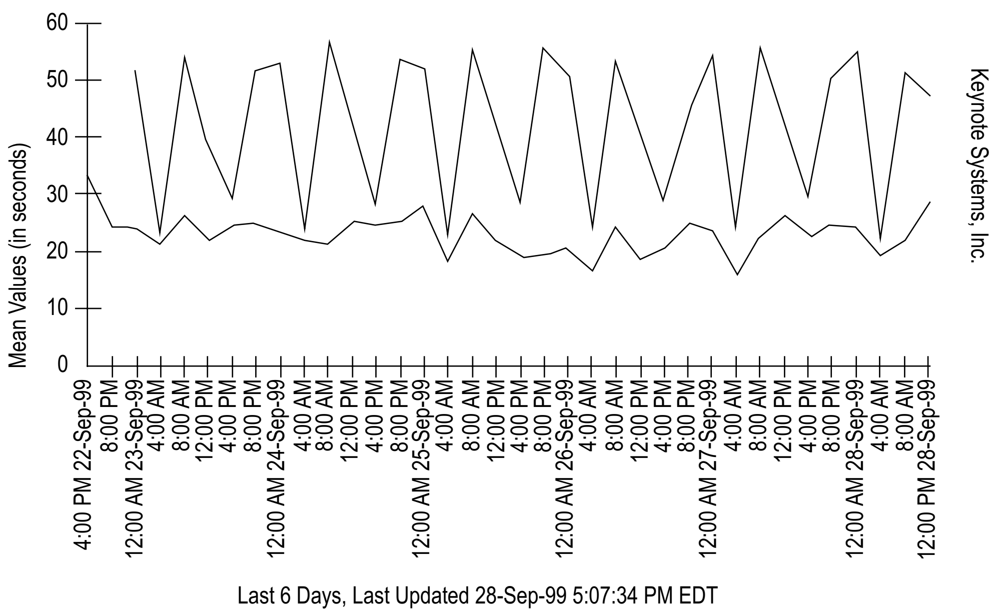


Figure 2: Web site performance by time history

to Contents, page 1

A governing principle in assigning users to user classes with different priorities is that the resources spent in analysing the request should be significantly lower than the cost of processing the request. Otherwise, the best-effort method is sufficient to provide service. Examples of rules determining user classes include origin and destination IP address, server port, URL, cookies, and encryption (it can almost be as expensive to decode an SSL request than to just process it). Request classification is done for each request and it provides the mechanism by which administrators can assign relative priorities to user requests (Fig. 1).

By using SLOs, the administrator can guarantee the order by which each request will be processed but not the level of resources assigned to specific applications. Service classes define the relative priority to access system resources for each application such as billing or credit card checking. Service classes are based on operating system tools such as *nice* on Linux or Process Resource Manager (PRM) on HP-UX, which allows the administrator to configure a user-defined scheduling policy for applications.

Commercial web sites may have a large number of replicated computer systems with one or more web server's instances per system. In addition to the concepts of web server instance and system, the administrator may want to define the concept of service as a set of servers residing on one or more systems. A server is defined by an IP address (to differentiate web servers residing on a single system with multiple LAN cards) and port. A server contains SLOs, which define the objective and associated corrective actions when the SLO is violated. Servers are also mapped to service classes to define the resource levels assigned to the server.

User classes are defined for the overall WebQoS environment. Each user is assigned to one of three classes (high, medium, and low) based on rules such as the user IP address. The flexibility of the architecture and the capabilities of redirecting HTTP requests facilitate the assignment of different levels of resources generic and registered users. Web applications with static and dynamic contents may also be assigned to different systems and platforms to optimize resource consumption at the server level.

At the overall e-commerce site level, the analyst defines the topology of the server farm. A server farm includes firewalls, application, and database servers, as well as the Intranet nodes, a proxy server for caching, a load-balancing server to distribute the transactions among the available servers according to the current load and resource availability or each server, and the servers under control of WebQoS.

Once the profile of the web server is defined, the administrator configures WebQoS using a Java-based graphical user interface (GUI). Since web server farms are intended to run non-stop, configuration is propagated into the appropriate systems. From the user's perspective, there is no perceived change when accessing the web server other than a faster response time and adequate throughput when the system is busy. For example, the chart shows the oscillating response time as measured by an independent company (keynote.com) on a standard web server and the more stable response time on a server that uses quality of service techniques to provide the same data (Fig. 2).

User requests differentiation allows the administrator to provide a uniform level of service to preferred users and opportunistic service to other users. This feature allows the customer to estimate the economic value of the WebQoS products for their particular web traffic patterns and e-commerce usage. If the system is not working at its peak capacity, all requests are processed immediately. When the system reaches its peak capacity, WebQoS also works as a service differentiator: higher-priority users have guaranteed access to platform resources while lower-priority users are deferred a predefined number of seconds, redirected to other sites with static-content pages, or rejected. That way, higher priority users have a better response time because their requests are processed sooner and lower priority users know earlier that the web server they are trying to reach is too busy to serve them.

The current implementation of WebQoS is architected in a client-server topology (Fig. 3).

- The WebQoS library intercepts the calls to the socket library issued by the web server so it can manage the request queues. When a request arrives to the system, WebQoS checks first if it is active for the particular port and IP address and, if it is, it classifies the request in one of the high, medium, or low priority classes and links it into the appropriate queue. The library also retrieves the next logical request from the queue using the scheduling mechanism defined by the administrator. The user may also invoke the

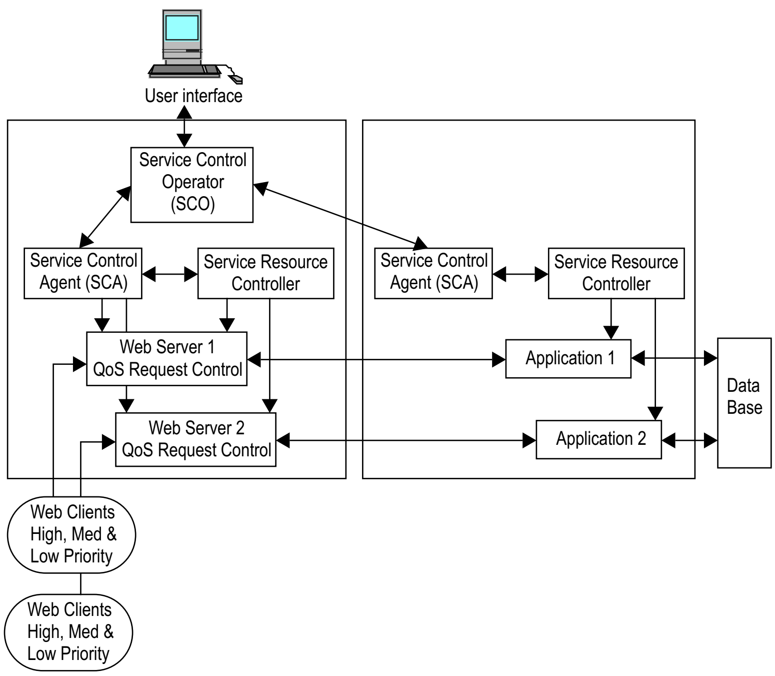


Figure 3: WebQoS architected in a client-server topology

to Contents, page 1

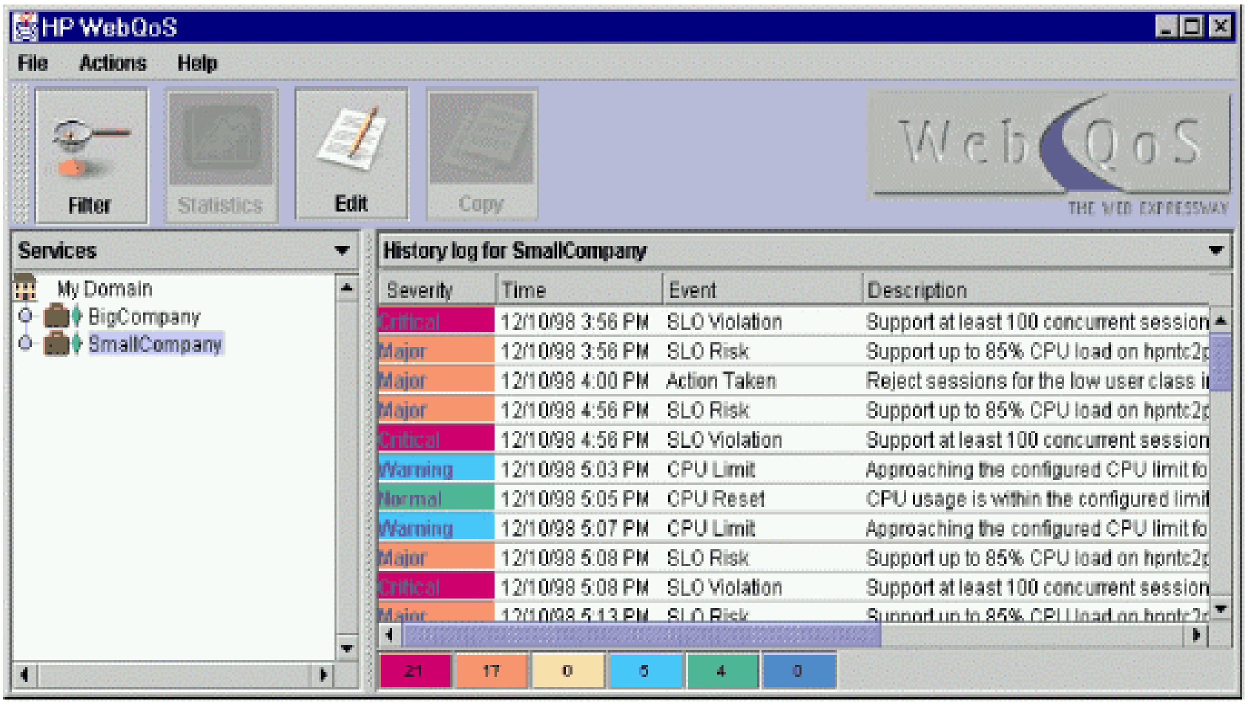


Figure 4: Screen describing the status of the SLOs

WebQoS API to assign application-defined priorities to requests.

- The service control agent (SCA) sits on the system where the web server is running and communicates with all the servers on the system to provide them the WebQoS parameters needed by the server and to gather all information and log requests sent by all the servers.
- The messages gathered by the SCA are sent to a central service control operator (SCO). The SCO talks to the agents (SCAs) to provide them information about WebQoS configuration and to request status data so the GUI can display the overall status of the domain and the specific status of services, systems, web servers (sites), and SLOs.
- The communications library is used for communication between the objects in the WebQoS domain. The protocol sits on top of TCP/IP and facilitates writing clients other than GUI to perform complex administration tasks.
- The tracing and logging daemon allows creating centralized repository to help the user understand the behaviour of WebQoS under the load specific to a WebQoS domain as well as to derive some data to analyse the access patterns from the users.
- The GUI is used to configure WebQoS and to analyse its behaviour. The figure shows the screen that describes the status of the SLOs (Fig. 4).

The architecture of WebQoS facilitates its integration with other tools. For example, while a graphical user interface is a simple and convenient tool for novice users who need to configure small networks, large web service providers may need automated tools to replicate a configuration change to hundreds or thousands of web server instances. Also, the history log for web server transactions may need to be correlated to other system or network-level events. Application servers may want to use the WebQoS library to extend the behaviour of the web server prioritization to all the steps of a user's request. Because of that flexibility, WebQoS core software and GUI interface is supported on a variety of web servers running on operating systems such as HP-UX, Linux, Windows NT, and Solaris.

References

[Becker/Geihs 98]
C. Becker and K. Geihs: Quality of Service. Aspects of Distributed Programs. International Workshop on Aspect-Oriented Programming at ICSE'98, Kyoto, Japan.

[Bhatti/Friedrich 99]
N. Bhatti and R. Friedrich: Web Server Support for Tiered Services. IEEE Network (September, 1999)

[Bouch/Sasse 99]
A. Bouch and M. A. Sasse: Network QoS. What do user need. Proceedings of IDC'99. Madrid, Spain September 1999.

[Cherkasova/Shankar 00]
L. Cherkasova and P. Shankar: Achieving Load Balancing and Efficient Memory Usage in a Web Hosting Service Cluster. Technical Report HPL-2000-27

[Pruyne 00]
J. Pruyne: Enabling QoS via Inception in Middleware. Technical Report HPL-2000-29

to Contents, page 1

A Tool for the Construction of Adaptive Web sites

Rosa Carro, Estrella Pulido, Pilar Rodríguez.

The TANGOW system, Task-based Adaptive learNer Guidance on the Web, allows for the creation of corporate Web sites that adapts the presentation of the corporate information to users with different profiles and different interests. Web Engineering is an area that has gained special relevance in the last years.

1 Motivation

In the last years all the technologies associated to Internet have experienced a great peak. The possibilities are increasing, because the number of accessible Web sites through Internet increases continually. In many cases, the sites that become public are of small size and, frequently, they correspond to individuals who create and maintain their pages Web in a handmade way. However, in the case of corporate Web sites, the creation and maintenance procedure cannot be improvised, since they seek to provide the users with a great quantity of information that, if it is not organised orderly, can lead to the degradation of the quality and integrity of the long term data.

To avoid the above problems, it is becoming increasingly evident that for the development of Web sites of a certain importance, more formal methods or specific computer tools should be used. The new techniques should certainly address the key required characteristics of these application types, in particular a) the necessity for rapid update of the available information, not only relative to a single item, but to the insertion of new sections and modification of existing sections, and b) the diversity of the profiles of the users interested in visiting the site. There are of course other functions demanding formalised control.

In response to the above perceived requirements, a new discipline emerged in 1997 in the field of the research and development of Web based systems: Web Engineering (WebE). The term was coined by San Murugesan of the University of Western Sydney Macarthur, Australia. An example of the growing interest in this field is the series of international workshops on WebE taking place every year since [Murugesan 98].

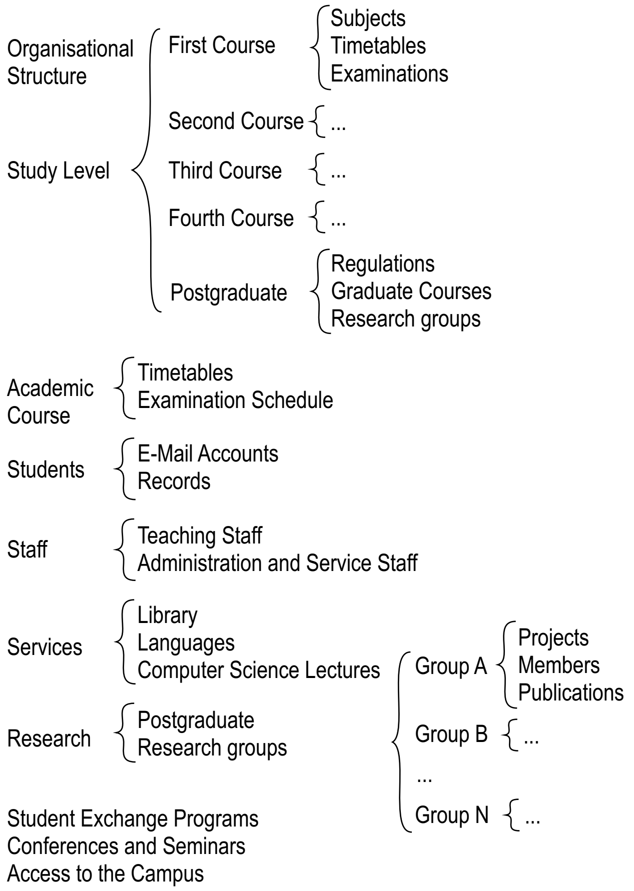


Fig. 1: Web site structure for a teacher

The objective of providing appropriate methods for the development of corporate Web sites urges the use of computer based tools that facilitate the creation as well as the maintenance of the information. There are plenty of tools though, such as Front Page [Jennett 00], for the creation and maintenance of HTML pages. But these tools provide no help in the administration and structure of the information. Neither do they assist in the Web site maintenance, and the necessary creation of views customised to each individual user's profile, taking into account the updating of the information to be included for each

Rosa María Carro Salas is assistant professor in the *Escuela Técnica Superior de Informática* of the Autonomous University of Madrid where she obtained the title of Informatics Engineer. Presently she is writing her doctoral thesis in the field of the dynamic generation of adaptive courses via Internet. Her line of research centres in the field of adaptive teaching, learning systems in the Web and multimedia in education.

Pilar Rodríguez is professor in the field of languages and systems at Autonomous University of Madrid. After graduating in physics at the *Universidad Complutense* of Madrid joined the IBM Scientific Centre in Madrid to take part in research projects in computer-assisted instruction and computational linguistic and other fields. She holds a PhD from *Universidad Complutense* Madrid. After a stay at Munich Technical University she joined the informatics engineering department of the Autonomous University of Madrid.

to Contents, page 1

profile, and the inclusion of new profiles. Most important are techniques that provide automatic adjustment of content and structure to best fit each user's profile. These techniques are part of the more general domain of adaptive hypertexts and hypermedia [Brusilovsky et al. 98], and it is in this line that the TANGOW system proposes itself as an adequate tool for the creation and the maintenance of corporate Web sites where the definition of distinct views in function of the users' profiles is essential.

2 Characteristics of TANGOW

The TANGOW system was originally designed to assist the development of adaptive Internet-based courses [Carro et al. 99b]. But it can also be used for the creation and maintenance of Web sites that adapt to the users' characteristics. In the case of a corporative Web site for a university, the user characteristics to consider could be, e.g., his language and the interests that led him to visit the Web site. These interests will be different if the user is, say, a student, a teacher, or a researcher.

The procedure used by TANGOW for generation of adaptive Web sites consists of designing for each user the contents that are to be presented to him, and the site structure which is specified as a group of content units. In TANGOW the term task is used to refer to such content units.

A content unit can be atomic or decompose to more specific content units. This user-dependent decomposition is rule-driven. The left part of the rule pertains to the compound content unit, while the units in those that breaks down appear in the right part of the rule. The process of decomposition, to which we will now refer as sequencing, can be of different types, but two sequencing types are of particular interest for corporate sites. The first is of type AND, which specifies that the user can access subunits in the order he prefers. If the subunits represent different options the user need not visit all, and the applicable sequencing is of type OR. Additional sequence types are of great utility when the system is used for educational ends.

Each rule has an activation condition containing the conditions related with the user's profile, or conditions related to the actions carried out by the user while interacting with the system; these conditions determine whether a rule is activated or not. This allows the definition of different forms of decomposition of a task in functions of the user's profile, by defining rules with the same task in the left part but different tasks and different activation conditions in the right part of the rule.

For the construction of contents, note that with TANGOW the pages are not built at design time, but are generated dynamically at execution time. The designer specifies the type of contents that should appear in the page and how they should be organized. The actual contents (hypermedia elements) are organized in agreement with characteristics such as the language they are written in, the level of difficulty, etc. At execution time, the system selects for each content unit the hypermedia elements associated with that unit and best adapted to the user's profile. The final page as shown to the user is obtained by a

Fig. 2: Examples of compound and atomic tasks (tarea)

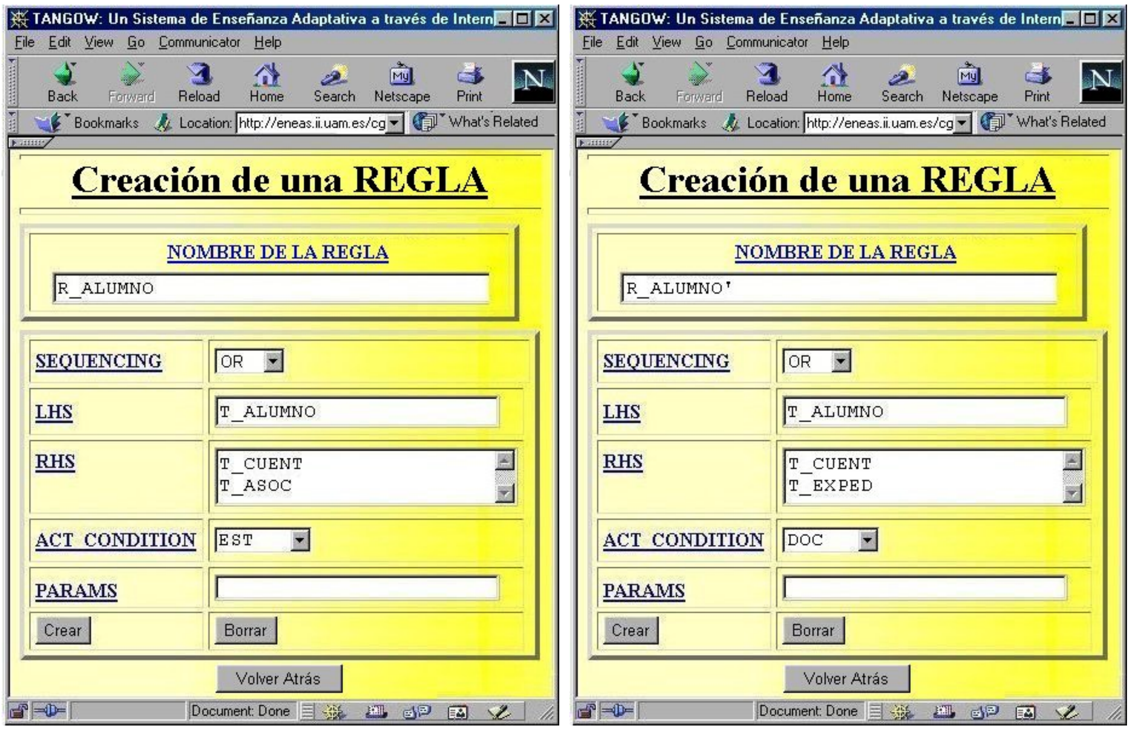


Fig. 3: Two different rules (regla) to describe the decomposition of the task

combination of the selected elements according to the pattern specified by the designer.

The structure of a Web site together with the contents are stored in databases. This greatly facilitates the maintenance tasks. Also, the independence of the structure of the Web site from the content of the pages presented to the user allow that aspects of the structure can be modified without having to change the content, and vice versa. On the other hand, the reuse of parts of the structure or of the content of a existing Web site when creating a new Web site is facilitated.

Another important feature of the system is that information is recorded about the user's interaction with the pages he accesses, how long he remains in each page, etc. This information can be used to reorganize the web site so that the most visited pages can be accessed more quickly. On the other hand, if many users spend very little time in a certain page this suggests that this page is just an intermediate step to arrive to where it is really wanted. This is usually annoying for the users who must click repeatedly, and it can be avoided by merging all the intermediate steps into one single page.

3 Design and maintenance of Web sites with TANGOW

When a designer creates a Web site, he defines the group of tasks and rules that constitute the structure or different possible structures of the Web site and associates these tasks with the hypermedia information elements and their disposition that will appear on the dynamically generated HTML pages. To illustrate this process, we will use the example of the design of a Web site with information about an university college. One of the possible structures of this Web site is the one shown in Fig. 1 that will be presented to navigators belonging to the profile "teacher".

The Web site generation tool provides an interface for the design of basic units or tasks, where the designer only introduces information about the task name, its description in several languages, the type (in a Web site it will usually be theory), and its atomicity. The other parameters are generated automatically by the tool, except for those that appear in the HTML field. In this case, the designer indicates the generic name of the hypermedia elements that will be used to compose the pages associated with that particular task.

Fig. 2 shows the use of the design tool to describe two tasks that are part of the college's Web site. The first task is of type theory T_ALUMNOS that includes information about the students. This task is compound, has no associated hypermedia elements, and will break down later into several subtasks. The second task, T_CUENT, is atomic and of type theory. The only hypermedia element associated with this task is identified as ST_ACCOUNT, and includes information about students' accounts.

Once all the basic units that are part of the Web site are defined, or during their definition, the designer introduces the rules that make up the structure of the Web site. As already mentioned, the Web site can be personalised by defining different rules that indicate different decompositions of the compound task, resulting in different structures for the same Web site. The activation of one or another rule will depend on whether the activation condition associated to each rule is satisfied, and it will determine the final structure of the Web site.

In our example, the designer has decided that the information accessible to the students will be different depending on the

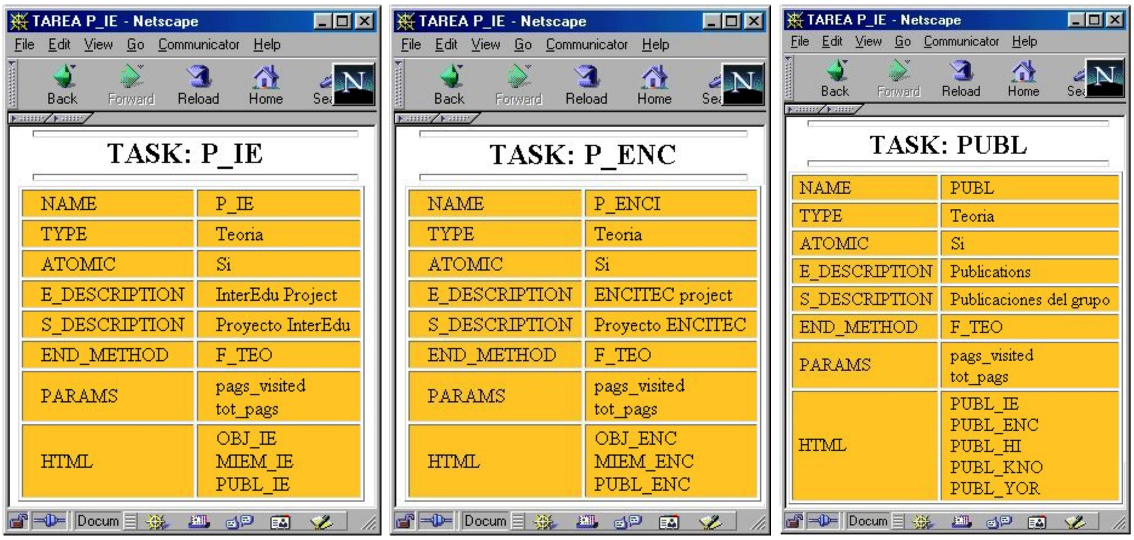


Fig. 4: Reuse of hypermedia elements

profile of the navigating user. If the user is a teacher, he will have access to the information about the students' accounts and records, as shown on Fig. 1. On the other hand, if the user is a student, he will have access to the information on his own student's account, and to the different students associations of the college, but he won't have access to his academic records. This adaptation to the visitor's profile is achieved by means of the definition of two rules defining the decomposition of the task T_ALUMNO (see Fig. 3). Each rule has a different condition of activation, so that the first will be activated if the navigator is a student, while the second will make it if it is a teacher.

The designer need not provide pages with menus to allow the navigating among the different parts of the Web site, since these are generated automatically by the system just before being presented to the students, and its content depends on the visitor's profile, allowing that the structure varies according to that profile.

Once all the tasks and rules that are part of the Web site are defined, the designer must define a main task, where every guest begins his visit. Next, the designer will click the button "compile Web site", which will check the consistency between the tasks and defined rules, and generate the instructions to store them the database of the Web site.

Finally, the designer must provide and classify the hypermedia elements that will appear in the pages generated dynamically during the user's interaction with the Web site. These elements will have the same name as the elements indicated in the HTML field of the tasks, and they will be classified in different directories, depending on their own characteristics and the profile of the potential visitors of the Web site (e.g. the language or the difficulty level). Thus, TANGOW chooses the most appropriate elements for each visitor in the moment to generate the HTML page associated to a task. An important application of this dynamic generation of pages is multiligualism.

Note that the fact that the pages are composed from several independent hypermedia elements allows the reuse of these elements in different tasks. This is the case shown in Fig. 4, where one can observe that the publications associated to the different projects of a research group (PUBL_IE, PUBL_ENC, etc.) are accessible not only in the HTML pages generated for each project that also include information about the objectives of the project and the members participating, but also is bundled in the page generated for the task PUBL, where all the publications of that research group is shown.

Another feature is that the designer can define a common style for all the pages, their background, header, and footer, so that the Web site will have an uniform aspect.

The maintenance of a TANGOW-generated Web site is very simple. To modify a hypermedia element, one simply edits and modifies it, without needing to access information on the Web site structure (tasks and rules). To replace an element by another, the old element is deleted and the same name is given to the new element.

To include new information in the Web site, the designer defines the tasks and rules that describe the new information and its structure, and relates these tasks with the information included previously in the Web site by means of some of the already existing rules or by defining some new rule.

Including a new profile is also a simple task for the designer using TANGOW. He must only define new rules for the access to information by the new profile, and as activation condition that the user belongs to that profile, together with other prerequisites in the rules associated with subtasks which were not accessible for this profile.

If the designer wants to modify the structure of the Web site, he modifies and/or deletes existing rules and tasks, and creates new rules and tasks as necessary.

If it is necessary to delete a portion of the Web site, the tool allows to simply carry out this operation selecting those rules and tasks that are to be eliminated.

In any case, when the designer presses the button "Compile Web site", the system checks the consistency of the Web site, checking if there are any task or group of tasks and rules referenced by any previous rule (defined but inaccessible from the main task), or if any rules reference to nonexistent tasks.

Also, neither is it necessary to update the links between the different pages of the Web site, because the pages that include links to other tasks are generated automatically from the updated information in the tasks and rules database, so that there will never be any connection to a missing task nor any missing connection to an accessible task.

4. Conclusions

The TANGOW system allows the generation of Web sites whose structure and contents adapt to the user's profile. Site structures are defined independently of the contents of the HTML pages that will be generated dynamically just before being presented to the users. Each structure is defined by means of a group of tasks that represent the basic units of the Web site, and their rules that define the relationships among the different tasks.

Also, the use of databases to store the hypermedia elements that will appear in the HTML pages, as the tasks and rules that are part of the Web site, facilitates its maintenance, and allows the efficient execution of the adaptation process. The consistency of the Web site is guaranteed, since the connections between the content units or tasks are verified automatically based on the information stored in the databases.

TANGOW is a design tool based on HTML forms, so that the designers can create and modify their Web sites easily. This tool will also soon include group working functionalities related to the cooperation among designers. These functionalities are based on the administration of Web sites and designers' identification, so that a designer may authorize or restrict access and modification rights on their Web sites to other designers.

References

[Murugesan 98]
WebE Home, http://fistserv.macarthur.uws.edu.au/san/

[Jennett 00]
Jennett, M.: Frontpage 2000 Developer's Guide (Professional Developer's Library). Osborne McGraw-Hill, 1999.

[Brusilovsky et al. 98]
Brusilovsky, P., Kobsa, A. and Vassileva J. (Eds.): Adaptive Hypertext and Hypermedia, Kluwer Academic Publ. Dordrecht, 1998.

[Carro et al. 99a]
Carro, R.M., Pulido, E., Rodríguez, P.: Task-based Adaptive learNer Guidance On the WWW: the TANGOW System. Second Workshop on Adaptive Systems and User Modelling on the Web at the Eighth International World Wide Web Conference, Proceedings online, Toronto, Canada, 1999.

[Carro et al. 99b]
Carro, R.M., Pulido, E., Rodríguez, P.: Dynamic Generation of Adaptive Internet-based Courses. Journal of Network and Computer Applications, 22, 249-257. Academic Press, 1999.

Acknowledgement

This work has been carried out in the project InterEdu, financed by the Spanish Interministerial Commission of Science and Technology (CICYT), I project number TEL97-0306.

to Contents, page 1

A Software Architecture for Search and Results Visualization on Intranets

Omar Alonso* and *Ricardo Baeza-Yates

In this paper we present a software architecture for search and results visualization on Intranets. Our approach is based on a model that defines an intermediate representation and a language for data interchange between the information retrieval software and the visualization software. Given that intermediate representation a designer can provide different visualization options for search and results visualization.

Keywords: Intranet, Architecture, Search, Information Visualisation

1 Introduction

The acceptance of World-Wide Web (WWW) technology allows access to existing systems in different systems using a common interface – a Web browser – from different geographical locations. Companies are trying to mimic Internet services and concepts in their Intranets.

There are two main WWW-based applications: database access and search and information retrieval. In the first one the goal is to migrate previous client-server applications to WWW as part of the eCommerce solution. Our focus will be in the second application: search and information retrieval [Baeza-Yates/Ribeiro-Neto 99], [Frakes/Baeza-Yates 92]. Our approach also explores the integration of different components and XML (eXtensible Markup Language) [Goldfarb/Prescod 98], [W3C 98] for data transport.

The process of information resource discovery on the Web is a primary task that requires essential tools. Lately there has been a growth in the number of tools and applications in the information retrieval area that is supposed to make the process easier for users. However searching the Web is still, sometimes, a frustrating experience. Not only because of the underlying search engines technology but also because of the current user interfaces.

Information retrieval and visualization are core components of more complex systems like digital libraries [ACM 95], [IEEE 96] and knowledge management systems[O'Leary 98].

Information access has two main components: 1) search and retrieve and 2) visualization (analysis and synthesis) of the result set.

The organization of the article is the following. Sections 1.1 and 1.2 describes briefly the information seeking process (search and visualization). Section two presents an Intranet search service as a case of study. Section three describes the model and software architecture of a search service for an Intranet. Section four enumerates implementation details. Finally we presents our conclusions along with a reference list.

Omar Alonso is a technical product manager in the *Databases and Application Servers* Group of Oracle Corp. His interests are information retrieval, digital libraries, visualization of information, and domain engineering. He graduated in Systems Engineering at the *Universidad del Centro* in Tandil (Argentina), Magister in Science from the *Universidad de Chile,* and Masters in Computer Science and Applications from Virginia Tech (USA). – The opinions expressed in this paper are exclusively his own and do no necessarily represent those of Oracle. <oalonso@us.oracle.com>

Ricardo Baeza-Yates is Ph.D. in Computer Science (University of Waterloo, Canada). Magister in Electrical Engineering, and Computer Science and Electrical Engineer of the *Universidad de Chile*. Presently he is Professor in the Computer Science Department of the *Universidad de Chile*. He is also president of the Latin American Centre for Studies in Informatics (CLEI), an association of more than 50 computer science departments in Latin America. His fields of research are information retrieval, algorithms, and data visualisation. <rbaeza@dcc.uchile.cl>

1.1 Information Retrieval

In the search process the user specifies in a certain way what he is looking for and the system presents the results set with a tabular representation. The results are not always what the user was interested in so the next step is to reformulate the query until the user finds (or not) the desired information.

The typical generic scenario for searching, retrieving, and displaying information on the Web is the following (see Figure 1). A user has an information need about a certain topic. With a user interface he formulates a query to the system (1). The query starts an action in the system (search engine, information retrieval system, digital library, or other software component) (2). The system will retrieve (or not) objects and will display them with appropriate messages and layouts in the same graphical user interface (GUI) where the user entered the query (3). Finally, the user decides if the documents are relevant or not. He can either exit the system because the information was found or refine the query and start again (4).

The process of searching information on the Web can be mainly performed in the following two ways:

- The user types keywords or phrases in search engines (e.g. AltaVista, Google, Hotbot, Northern Light, etc.) or browses and also searches over predefined classifications (e.g. Yahoo!). The technology used here belongs to the domain of

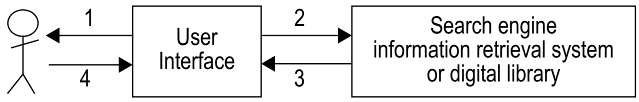


Figure 1: The information seeking process

information retrieval. The underlying technology of a search engine, as an example, can be simple inverted files, vector space models, or Boolean systems [Baeza-Yates/Ribeiro-Neto 99], [Frakes/Baeza-Yates 92]. Harvest is also a well known architecture for Web searching. [Browman et al. 94].

- Using an autonomous software agent to perform searches and inform the user of the results. This type of agents belongs to the area of artificial intelligence or agent-oriented programming [Bradshaw 97]. A very useful form of software agents is as an information retrieval agent, whose main function is to find and gather relevant information [IEEE 97], [Jacobs/Shea 96].

In an Intranet environment there is much more control over content that on the Internet. By definition the Web has little structure, but it is possible to have more control over Intranet than Internet sites.

1.2 Information Visualization

Current search engines interfaces are intuitive and in some cases restricted by the nature of the Web. There is limited use of colour, the interaction is minimal and several common desktop features are inexistent. Although the trend is changing, the current search interface is a text box for the input query and a search button.

The process of visualizing results set can be pure textual or more visually appealing with a visualization metaphor. With a pure textual result, the user has to read through an ordered list of ten or twenty items (usually documents) that contains the desired information. This is acceptable if the answer is in the first page, otherwise the user has to examine all the entries in the list. Usually the data about an item are the abstract of the document, the URL (Uniform Resource Locator), title, last modified date, and size. All this data is compressed into three or four lines. With a click on the URL the user can open a document.

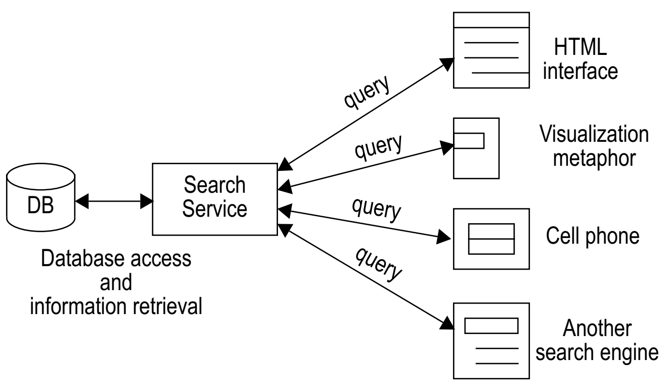


Figure 2: A search service

The user usually can not manipulate how to display the result set on the screen and he can not change the interface at all. This approach works if the items to be found are in the first couple of Web pages. Otherwise it is a tedious process until the user has found what he was looking for. A graphical visualization metaphor solves this problem with a richer interface where the user can query and navigate the result set on the same page.

Information visualization is an effective tool for partially solving data overloading problems in Web information retrieval [Stuart et al. 99], [Hearst 99]. Another approach for exploring WWW search interfaces can be found in [Greenberg/Garber 99]. Many interfaces are fixed and do not allow tools like plug ins to improve them.

From the hardware point of view, the process of searching and retrieving information in WWW is limited to certain high resolution displays. Although there is a wide range of devices for the Web, a sophisticated visualization metaphor is not practical. Our approach provides a solution to this problem. The user can select a visualization metaphor for a particular hardware device if its available on the system.

2 Search and Results Visualization in an Intranet

Let's see a practical Intranet application in a given company. The problem be the following: the user wants to search information using a given software and wants to see the search results using a different visualization tool. For example, an engineer may use his favourite browser as the main interface for search and visualization. But when he is travelling, he has no access to his high end workstation to search for information. Maybe he has another type of device to access the network like a personal assistant (e.g. Newton or PalmPilot) or cell phone[1].

We will describe both aspects: search and visualization.

2.1. Search

The scenario be the following. The company maintains a centralized search service which controls access to main servers and also to several inside web sites. Different users with different hardware devices can access the service and perform search queries. The system also has a subscription feature and a personalization mechanism.

Figure 2 shows an instance of the search service with some clients that are performing search queries.

The search service provides an API (Application Program Interface) that allows interaction from different visualization metaphors. Without giving any particular details of a programming language, a call to the search service has the following syntax:

```
search -query -id -collection -results
```

The call to the search service is one code line (`search`) with four parameters: the query (`query`), client or customer identification number (`id`), collection (`collection`), and result set (`results`). As an example and following the previous figure, a group of users are issuing the same query: "information visual-

1. We assume that there is some sort of authentication.

ization". A user (id=12) will prefer an HTML interface from a web browser. The web interface will make the following call:

```
search -query 'information visualization'
-id 12 -collection 'intra' -results 1
```

In another example, the same user would like to view the results using a visualization metaphor like HBP [Baeza-Yates 96]. HBP's Java code will execute a JDBC connection to call to:

```
search -query 'information visualization'
-id 12 -collection 'intra' -results 2
```

Another user (id=20) issues a query from his cell phone where the interface makes a call to:

```
search -query 'information visualization'
-id 20 -collection 'intra' -results 11
```

Finally, a parasite search engine uses the search service to provide answers to a given query:

```
search -query 'information visualization'
-id 2 -collection 'para' -results 50
```

The identification mechanism is very important because it allows a level of personalization in the application. In this case the search service can retrieve a profile of the user every time he is opening a connection to the service.

The result set of the search query is a list of items usually ordered by relevance.

2.2 Visualization

Let us consider another point of view and imagine a user who has learned to use an interface and is happy with it because it works. We will not give implementation details but we assume it is complex and the user considers it a valuable tool. The user wants to operates this metaphor as the main way to access any kind of information at any time.

The user performs queries against different information sources using the same visualization metaphor (Figure 3). This can be done if the sources share the same data representation otherwise this level of integration is very difficult.

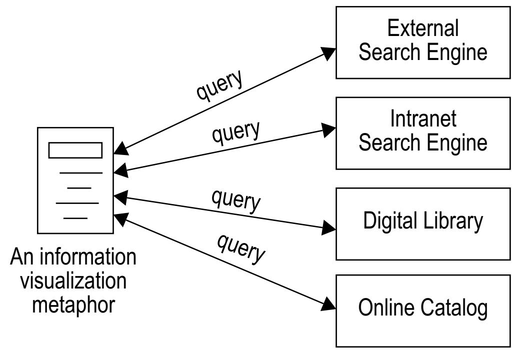


Figure 3: Visualization using different sources.

3 Model and Software Architecture

We want to make the search component independent from the visualization component in a standard way. By standard we mean that different visualization metaphors can interact with different retrieval software and vice versa. The user interface is independent from the retrieval software and at some point is user defined from a set of available interfaces [Alonso/Baeza-Yates 98].

Conceptually our model consists of three main parts: a set of searchers, a set of visualizers, and a markup language that acts like an independent representation that allows both sets to interchange data.

Figure 4 shows the user's perspective with the system in action. A circle represents a searcher. In the example, B_1 is a software agent with information retrieval capabilities. B_2 is an specific application that contains an information retrieval component (assisted by a crawler) and that also access the operating system and a database. A thick circle represents a visualizer. For example, V_1 is an implementation of HBP [Baeza-Yates 96] and V_2 is a hyperbolic browser [Lamping/Rao 96].

Searchers and visualizers interchange information using an intermediate representation that we will introduce later. There are different ways to manage this representation in the implementation. For notation purposes only we suggest that the choice is to store the data in a temporary files of type RI (intermediate representation).

A user has an information need and define his searcher (B_1) and visualizer (V_1). The searcher B_1 performs a search over a WWW source and we can define it as a search engine. The visualizer V_1 is an implementation of HBP or a visual metaphor similar to TileBars [Hearst 95]. B_1 produces as output an intermediate format (RI). V_1 reads this file and displays to the user all the search results using a visualization metaphor.

Another user specifies a different searcher (B_2), and visualizer (V_2). The searcher B_2, performs a different type of search using a complex information retrieval software that combines different data sources (Web, databases, and flat files). B_2 produces output to a file

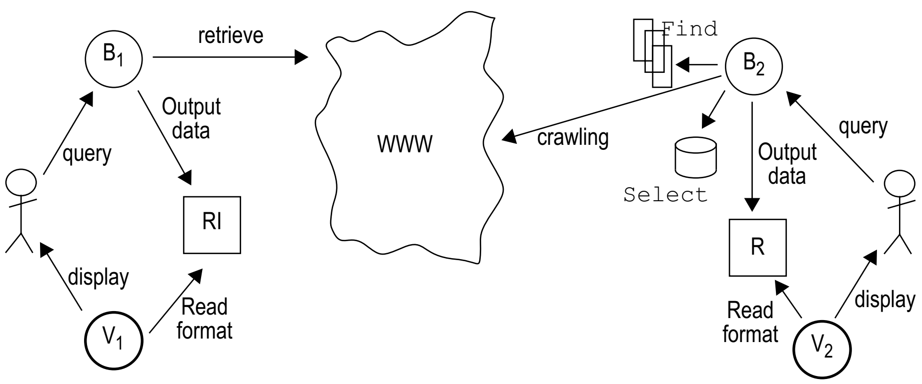


Figure 4: A General Description of the Model and its Components

(RI) using the same mechanism as the prior examples. V_2 is an implementation of the hyperbolic browser HBP and will display the result set.

3.1 Intermediate Representation

The search and visualization components need an intermediate representation for data exchange. We are interested in the following results data from a search query. A visualization component will later read this data for presenting the results to the user.

- Query
- URL
- Title
- Abstract
- Author
- Content-length
- Content-type
- Last-modified
- Score or rank

This intermediate representation will be used by the search and visualization components. If a component (search or visualization) does not understand the format, a transformational engine will apply a set of rules to transform the intermediate representation to its internal format. The following is the search query results for "information visualization". The output from a search component will be an ordered list of items each of them containing: rank, URL, abstract, last modified data and page size (among others).

```
1
http://www.cs.panam.edu/info_vis/home-info_vis.html
Information Visualization at U. Tx. - Pan American
Information Visualization. at University of Texas
- Pan American. The purpose of
computing is insight, not numbers.
- Richard Hamming, 1962. The goal of….
Last modified 18-Oct-96
page size 4K
```

```
2
http://www.elastictech.com/
Elastic Technology - Information visualization java tools
Java technology to visualize, navigate, and manage large information…
Last modified 27-Apr-99
page size 4K
```

```
3
http://graphics.stanford.edu/courses/cs348c-96-fall/resources.html
Some Information Visualization Resources on the Web
Information Visualization Resources on the Web.
This page is a partial collection of
online InfoVis resources. See the accompanying
bibliography for…
Last modified 29-May-99
page size 14K
```

We can see that all this information could be structurally and semantically richer if it can be expressed in a language.

The visualizer reads the result set as input and renders the information according to it visual metaphor. An advantage of this intermediate format is that the searcher can include different document attributes that can be (or not) used by the visualizers. On the other hand certain visualizers can prefer a searcher that provides more better quality data

3.1.1 IVL

We propose a markup language IVL (Information Visualization Language) as the implementation of the intermediate representation introduced in the previous section.

There are other intermediate representation formats in research and industry. A well know format in knowledge representation is KIF (Knowledge Interchange Format) used in artificial intelligence projects. In industry, MCF (Meta Content Framework [Guha]) was introduced some years ago by Apple. One of the most interesting applications of MCF is HotSouce. HotSouce is a 3D visualization metaphor for navigating Web sites.

Lately there has been much interest in extraction of semi-structured data from Web pages. The most important example is VDBMS (Virtual Data Base Management System) by Junglee [Rajaraman et al. 98] that makes Web and other external data sources look like a unique relational database. From a research perspective an interesting approach of extraction by example is presented in [Ribeiro-Neto et al. 99].

Our interest in a language is because we want to make the representation independent from any proprietary format. Also we can automate wrapper generation as in several semi-structure data extraction projects.

We are interested in XML and RDF [W3C] as the main options for implementing IVL in a Web environment. Why XML or RDF? With the current trend in eCommerce infrastructure, many formats (like EDI) are re-evaluated in the XML context. We believe that XML and RDF will evolve with Web technology and this will allow improvements on IVL. Today there are more products and tools around XML than RDF.

A query result set using an XML implementation is:

```
<ivl>
<rank> 1.</rank>
<url>http://www.cs.panam.edu/info_vis/home-info_vis.html</url>
<title> Information Visualization at U. Tx. - Pan American </title>
<abstract>
Information Visualization. at University of Texas - Pan American. The purpose of computing is insight, not numbers. - Richard Hamming, 1962. The goal of… </abstract>
<lm>Last modified 18-Oct-96 </lm>
<size> page size 4K </size>
```

to Contents, page 1

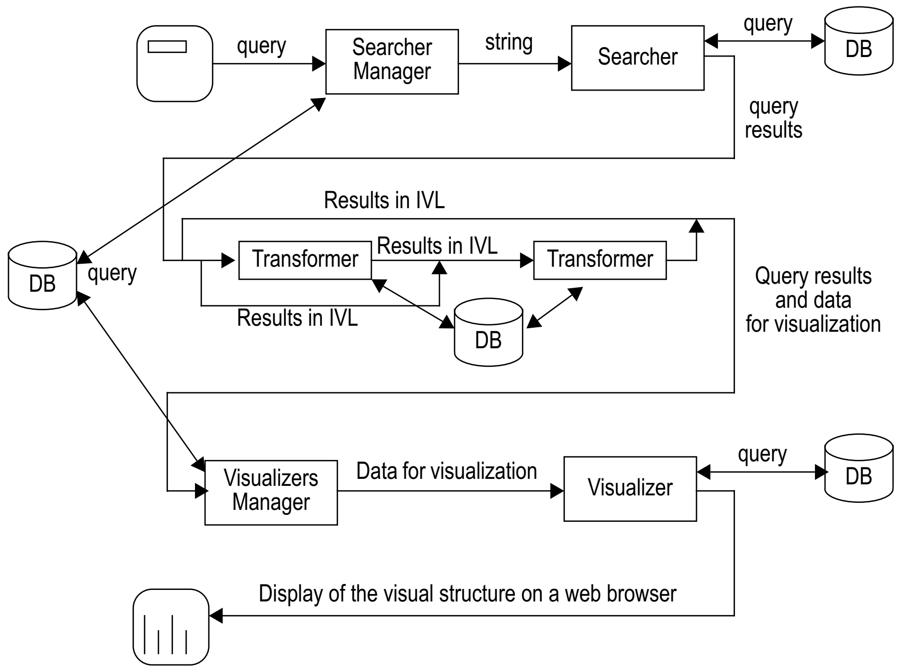


Figure 5. Software Architecture

```
<rank> 2.</rank>
<url>  http://www.elastictech.com/ </url>
<title> Elastic Technology - Information
visualization java tools  </title>
<abstract>
Java technology to visualize, navigate, and
manage large information…
</abstract>
<lm>
Last modified 27-Apr-99</lm><size> - page size
4K </size>

<rank> 3.</rank>
<url>  http://graphics.stanford.edu/
courses/cs348c-96-fall/resources.html</url>
<title> Some Information Visualization
Resources on the Web  </title>
<abstract>Information Visualization Resources
on the Web. This page is a partial collection
of online InfoVis resources. See the
accompanying bibliography for…
</abstract>
<lm>Last modified 29-May-99 </lm><size> page
size 14K </size>
</ivl>
```

3.2 Software Architecture

This section describes our proposed software architecture using a hybrid architectural style notation of Shaw and Garlan with repositories [Shaw/Garlan 96]. Figure 5 shows the architecture for our proposed solution.

The user has the choice of selecting a searcher and a visualizer as part of his profile. Then he issues a query to the system. The Search Manager then invokes the specified searcher to perform the query. Most searchers use some kind of data storage that we represented in the architecture as a data base, but it could be a file system, or some other structure. The results from the query can be in IVL format or not, depending whether the searcher supports IVL. If the searcher does not support IVL, a transformer will map the internal format of the searcher to IVL.

The process of displaying the results using an information visualization metaphor or a different user interface is the following. The Visualizer Manager invokes the specified visualizer that will display the results for the user. If the visualizer supports IVL then it just displays the results to the user. Otherwise there has to be another transformation from IVL to the internal format of the visualizer. The transformer also uses data storage (in this case a data base) for storing the transformed format.

It is important to note that the architecture describes components without specifying a particular vendor or a particular technology. These components work in a Web environment and our interest is to build software using them. The Open Source community is starting to identify this way of development. We will discuss technology issues in the prototypes section.

4 Implementation

In this section we describe at high level some implementation details. We call WWW-based computing as the integration of different heterogeneous data sources under one environment. We introduce two common models: CGI (Common Gateway Interface) and SSWA (Server-Side Web Applications).

In the first model, to develop a Web interface for an existing database, the programmer writes a CGI script using Perl, C or any UNIX shell script. This script is an external program run by the Web server to access the data and create HTML output. This HTML is then presented on the Web browser. This architecture works and has some know limitations like the difficulty to maintain state and session connection and extensibility.

The second model consists in adding HTTP functionality to the database software. Using JDBC with Java servlets to replace CGI is a better architecture that solves previous client-server problems and allows the server-side applications to interact with a wide range of databases. A Java servlet is like an

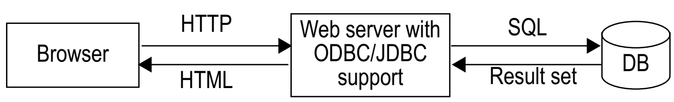


Figure 6: ODBC/JDBC-based architecture

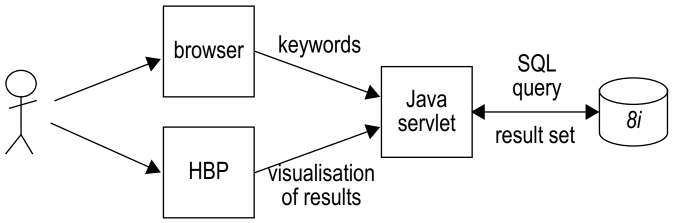


Figure 7: Search service architecture

applet except that runs in the server. Figure 6 shows an architecture based on ODBC/JDBC.

The search service prototype consists of interMedia Text (Oracle) as the underlying search technology and a simplified version of the horizontal bookpile metaphor as the visualization metaphor.

There are several ways of implementing the prototypes in a Web environment using Java technology: servlets, applets, JSP (Java Server Pages), and EJB (Enterprise Java Beans). There are some techniques about where to place the components, based on the degree of separation between presentation and business logic. Our emphasis is moving most of the computation to the server side so we choose servlets as the basic model. However we believe that similar development can be done using other server models.

The user issues a query from the browser. A Java servlet is responsible for opening a JDBC connection to the *Oracle 8i* database, issuing the text query, and retrieving the rows. With the XML-supported features from the product, the horizontal bookpile metaphor (HBP) reads the IVL format and displays the results to the user.

The servlet has to "massage" the results from that query and produce IVL markup code. HBP reads IVL and produce a graphical representation of the result set.

Figure 7 shows the architecture of the prototype. The back end side is an *Oracle 8i* database with interMedia Text and a Java servlet. The front end side is a variation of the HBP.

The main advantage of this solution is that has an industrial strength back-end that allows the prototyping of different visualization metaphors and user interfaces. On the other hand it requires a level of expertise in a number of components (database, interMedia Text, servlets, XML, etc.).

5 Conclusions

We have presented a model and a software architecture for searching and visualizing answers in WWW retrieval with emphasis in the visualization process. With so many systems and tools for WWW retrieval, our model defines a way to share the output of those so that the visualization is more independent and therefore an option of several to the user.

We believe that our architecture is suitable for current commercial Intranets. Our software architecture could be used to implement knowledge management solutions where information is in different formats in different places. The model can also be used to implement previews and overviews interfaces for digital libraries.

References

[ACM 95] ACM Special Issue on Digital Libraries (1995).

[Alonso/Baeza-Yates 98] Omar Alonso and Ricardo Baeza-Yates, "Visualization of large answers in WWW", *Proceedings of SSSC 98*, Antofagasta, Chile (1998). pp. 2–7.

[Baeza-Yates 96] Ricardo Baeza-Yates, "Visualization of Large Answers in Text Databases", *AVI'96*, Giubbo, Italy (1996). pp 101–107.

[Baeza-Yates/Ribeiro-Neto 99] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley Longman, Endinburgh Gate, London (1999).

[Bradshaw 97] J. Bradshaw (Ed.), *Software Agents*. The MIT Press, Cambridge, MA. (1997).

[Browman et al. 94] C. Browman, Peter Danzig, Darren Hardy, Udi Manber, and Michael Schwartz, "The Harvest information discovery access system", in Proc. *2nd Int. WWW Conference* (October 1994). pp 763–771.

[Stuart et al. 99] Stuart Card, Jock Mackinlay, and Ben Shneiderman. *Readings in Information Visualization*, Morgan Kaufmann, San Francisco, CA (1999).

[Frakes/Baeza-Yates 92] William Frakes and Ricardo Baeza-Yates, *Information Retrieval. Data Structures and Algorithms*, Prentice-Hall, Englewood Cliffs, NJ. (1992).

[Goldfarb/Prescod 98] Charles Goldfarb and Paul Prescod, *The XML Handbook*. Prentice-Hall, Upper Saddle River, NJ. (1998).

[Greenberg/Garber 99] Ilan Greenberg and Lee Garber, "Searching for New Search Technologies", *IEEE Computer*, Vol. 32, No. 8 (August 1999).

[Guha] R.V. Guha, "Meta Content Framework" (http://www.xspace.net/hotsauce/mcf.html).

[Hearst 95] Marti Hearst, "TileBars: Visualization of term Distribution Information in Full Text Information Access", *Proceedings of CHI'95*, Denver, CO (May 1995).

[Hearst 99] Marti Hearst, "User Interfaces and Visualization" in *Modern Information Retrieval* (by R. Baeza-Yates and B. Ribeiro-Neto), Addison-Wesley, Endinburgh Gate, London. (1999).

[IEEE 96] IEEE Special Issue on Digital Libraries (1996).

[IEEE 97] IEEE Internet Computing, Special Issue on Software Agents, (July/August 1997).

[Jacobs/Shea 96] N. Jacobs and R. Shea, "The Role of Java in InfoSleuth: Agent-based Exploitation of Heterogeneous Information Resources", MCC Technical Report #MCC-INSL-018-96 (1996).

[Jenkins et al. 98] Charlotte Jenkins, Mike Jackson, Peter Burden, and Jon Wallis. "Searching the World-Wide Web: An evaluation of Available Tools and Methodologies". *Information and Software Technology*, Vol. 39, No. 14–15 (February 1998). pp 985–994.

[Lamping/Rao 96] J. Lamping and R. Rao, "The Hyperbolic Browser: A Focus + Context Technique for Visualizing Large Hierarchies", *Journal of Visual Languages and Computing*, 7(1), (1996).

[O'Leary 98] Daniel O'Leary, "Enterprise Knowledge Management", *IEEE Computer*, Vol. 31, No. 3 (March 1998).

[Rajaraman et al. 98] Anand Rajaraman, STS Prasad, and Debra Knodel, "Junglee Virtual DBMS", in *The XML Handbook*. Prentice-Hall, Upper Saddle River, NJ. (1998).

[Ribeiro-Neto et al. 99] Berthier Ribeiro-Neto, Alberto Laender, and Altigran da Silva, "Extracting Semi-Structured Data Through Examples", *Eighth ACM Int. Conference on Knowledge and Information Management* (ACM CIKM'99), Kansas City, Missouri, 1999.

[Shaw/Garlan 96] Mary Shaw and David Garlan, *Software Architecture*, Prentice-Hall, Upper Saddle River, NJ. (1996)

[W3C 98] W3C, "Extensible Markup Language (XML)" http://www.w3.org/TR/1998/REC-xml-19980210

[W3C] W3C, "Resource Description Format" http://www.w3.org/TR/PR-rdf-schema/

to Contents, page 1

www.admin.ch – Access to the Swiss Parliament

Elisabeth Bürki Gyger and Nicolas Kessler

www.admin.ch is a information site of the Swiss Federal authorities serving as gateway to the Swiss Government and Administration. Since the launch of the Web site in September, 1995, it has been constantly developed and adapted to the needs of Internet clients.

Responsibility

www.admin.ch is the responsibility of the Information and Press service of the Federal Chancellery. The presence on the Internet falls under the category of information and makes up part of the mandate to inform the Parliament, the cantons and the public, as set down in article 10 of the Law on Government and Administration.

Overview

Information from the federal departments (ministries) and federal offices can be accessed directly from the home page. Along with information about the Government, the Administration and their activities, the site provides details about political rights (people's initiatives, referenda, public votes and national elections) and legislation in Switzerland. Of particular note is the systematic collection of laws.

The target audience includes citizens, Swiss nationals living abroad, the media, politicians, teachers and schools, as well as foreigners interested in Switzerland.

www.admin.ch reflects the multilingual nature of Switzerland with information in the three official languages German, French and Italian. Present imbalances in the amount of information provided in each language will be corrected. Wherever possible, consideration is given to the fourth national language, Romansh, and with the interests of an international public in mind, documents in English are appearing more and more. The diversity of languages is evident on each page: The language selector displayed on top of each page facilitates the search for information and is highly appreciated. The naming of the Web site was also influenced by this multilingual characteristic. The domain name was chosen particularly because it applies in all four national languages and English.

www.admin.ch	*Admin*istration	German
	*Admin*istration	French
	*Ammini*strazione	Italian
	*Admin*istraziun	Romantsh
	*Admin*istration	Englisch

The latin "**C**onfoederatio **H**elvetica" (Swiss Confederation), together with the Swiss Cross, was chosen for the logo. "CH" is the official (ISO) abbreviation for Switzerland.

How the site developed

The Federal Chancellery made the first Web trials in 1993. Thanks to the basic Web standards developed by Tim Berners-Lee at CERN – HTTP, HTML and URL – we could gather information from the most varied sources and make it available to a wide circle of people using the Administration's internal network. One of the prerequisites for the Web, the network protocol TCP/IP, was already being widely used in the Administration. Particularly in cases of "exceptional situations", where information has to be prepared in a short time without a lot of technical arrangements, related to other data and presented concisely, the Web appears to be the ideal solution.

To a certain extend we early users of what is known today as "Intranet-based knowledge management".

For more than a year, a lot of persuasion had to be used before more internal data were made suitable for the Web. "Surfing" was popular but many people shied away from the effort involved in preparing their "own" information for the Web. Nevertheless, the Federal Chancellery together with the Parliamentary Services, the Parliamentary staff office, appeared on the Internet with a vast array of information September, 1995 and thus became one of the first government representatives on the Web. Then, as today, what was of primary importance was not the presentation of the individual office, but rather substantial information about legislation and direct democracy.

Elisabeth Bürki Gyger is responsible for the management, the updating and the further development of www.admin.ch as well as for the coordination of the internet activities of the Swiss Federal Chancellery. She chairs the forum of the persons in charge of WWW in the Departments, the Parliament Services and the Federal Chancellery. Her domains of interest: Web design, especially graphics, text, and navigation design.

Nicolas Kessler, head of Information Systems, is responsible for the planning and execution of projects in the area of information modelling and retrieval. His/its interests include XML, topic maps, knowledge management, and the application of modern information technologies in decision processes.

to Contents, page 1

The diversity of design and structure of the Web sites when surfing through them is impressive. The decentralised management of the sites arises from the fact that many federal offices were on the Internet before their relevant ministries. This individuality managed to avoid creating the impression of an administrative monolith on the Web.

In this context, it is worth noting that in contrast to other countries the press and information services are decentralised, i.e. are located in the different federal ministries and offices. There is no central information and press office in Switzerland.

The diversity of the Web sites is however not particularly worrisome. In fact it has advantages. The distribution of the responsibilities has created room for manoeuvre in the ministries and offices. It gives them freedom to present complete information directly to their public thus guaranteeing the freshness and continuity of the information.

Today, information about the collection of laws and the legislative proceedings make up the largest part of the Confederation's Internet offer and records the highest number of "hits". For internal information systems and processes, suitability for the Web was the main concern. All header data and texts in their updated versions are managed with intranet based technology and constantly made available on the Internet in consolidated form. Whenever possible, cross references are included, pointing for example from the header data of a law put to a public vote to the detailed results.

Principles for the organisation of www.admin.ch

The development principles set down in 1995 are still valid for the people responsible for www.admin.ch. In fact some have even evolved into a kind of doctrine.

- A high degree of substantial information as well as a straightforward, easy-to-follow organisation of the individual pages. Access to politics should be made easier. Fears should be reduced.
- Access to information should be available to all, regardless of computer platform or browser software. Information should be accessible with older computers, small screens or slow modems. Consequently, new features are introduced only cautiously. Purposely, and possibly at the expense of visual attractiveness, elements that hinder or slow access to information are avoided.
- For Web graphics, attention is paid to keep memory volume as low as possible.
- Web pages must be accessible to the blind and visually handicapped (speech browser, Braille characters).
- The offer must permit future adaptations and expansions.

However, to maintain a minimum of coordination, a Web Forum was created in Spring 1999. In this group, departmental webmasters meet under the chairmanship of the Federal Chancellery. At present, work is focused on the revision of Web guidelines set in 1996. In future, information (directions, recommendations, tips) will be available on the Intranet for all Webmasters of the Federal Administration.

Prospects

The possibilities of the Internet as an information and communication medium are far from being exhausted and they will grow with the development of new technologies. Navigation improvements are planned for the near future. Topic maps will enable thematic access to information. At the same time, an alphabetical index will be available. In this connection considerable importance will be placed on terminology. A common language denominator will have to be found (example: drug = narcotic, addictive substance).

Another major information project concerns better reporting of the sessions of the Federal Council (government) and information on government activities in general. Studies are being made into the possibility of a live transmissions of government news conferences, as is already the case with parliamentary debates (www.parlament.ch).

Furthermore, the existing offer will be revised and Italian put on a par with German and French. Existing categories will be made even more informative and user-friendly. This means a renovation of the home page www.admin.ch. Under the catchword *e-government* new possibilities are being discussed relating to electronic communication with the government authorities and new forms of political participation.

The Spanish Administration on the Web

Gloria Nistal

The enormous speed in the diffusion of the Internet forces companies, citizens and society to change their habits in connection with information access. The Spanish administration could not distance itself from this authentic revolution. Aware that the web is a means of inestimable value for communicating with its citizens, the Spanish public service is establishing the base for the creation of a public administration portal that is evolving from the present administrative "information hypercentre" to an authentic "administrative information and processing hypercentre" by entirely electronic means.

The Internet is an easy to use and affordable technology for the great majority. This has caused its own evolution at an enormous speed. Internet users are the millions of people who want to do quickly what previously only happened in the imagination of experts. Internet is a sociological phenomenon. It is the clearest example for a disruptive technology, according to Schumpter's principle in which the political, economic and social models change linearly while the technology changes exponentially.

The real world happens in real time [Savetz 98] and the users want to get in real time the tools the technology offers them to improve their lives. For companies, it is important to recognise that the Internet is no longer a secondary, but primary way of making business. According Negroponte [Negroponte 99], we have passed from the atoms to the bits; in the same way electronic commerce, instead of being based on the physical movement of goods or paper, is based on the movement of entities of information like orders, electronic invoices or the Web pages.

This electronic bit flow occurs through digital instead of through physical channels. And that channel can be used equally as a marketing channel, as a distribution channel and as a communication channel, eliminating the considerable and expensive intermediate chain.

The current situation

A recent study by the OECD [OECD 00], and the last General Study of Media, highlights the important and rapid growth of the access to Internet in Spain, where there are already about 4 million Internet users, that is to say, 10% of the Spanish population. The Spanish public services, aware of the importance of the use of the electronic, informatic and telematic media[1] – and of the Internet in particular – for the improvement of their services and their communications with citizens, undertook some years ago a policy of diffusion of information and of electronic processing through the Internet.

The Ministry of Public Administration (Ministerio de Administraciones Publicas, MAP) created a web site http://www.map.es/ that includes conventional information about its structure, its organisation chart, and a section with government news, a section about new government projects, a search engine and an electronic mailbox for comments and observations. It also provides links to the National Institute of Public Administration, or to the civil servants society. Within this page the citizen find important links to Web sites we will comment on below in this paper, since they are the most visited pages of the Ministry.

According to the statistics for March 2000, the MAP Web site is among the three most visited sites of the Spanish State's General Administration. The three most visited links from inside this site point to:

1 the site of the Administrative Information Centre http://www.igsap.map.es/cia.htm. This centre contains useful information, including:
- The basic legislation of the General Administration of the State, of the Autonomous Regions and of the European Union
- Awards, scholarships and grants
- Employment offers in the public sector
- Location of the different administrative information centres of the public administrations
- Official bulletins of the State, the Autonomous Regions and the provinces
- Appeals and complaints before the Administration

Gloria Nistal Rosique is government employee of the *Cuerpo Superior de Sistemas y Tecnologías de la Información*. She is technical project leader of the Unique Window (*Ventanilla Única*) project and of the pilot project *PISTA-Ventanilla Unica*. She is member of the Monitoring Committee of of the Metric Methodology v.3., head of the Software Approval Group of the superior computer science Council, and head of the Administrative Information Hypercentre (HIA). She is Vice-president of the ATI Madrid Chapter (*Asociación de Técnicos de Informática*) and Coordinator of the of of the Software Quality Group of this Chapter, as well as Vice-president of the Software Section of the Spanish Quality Association (1994-1998).

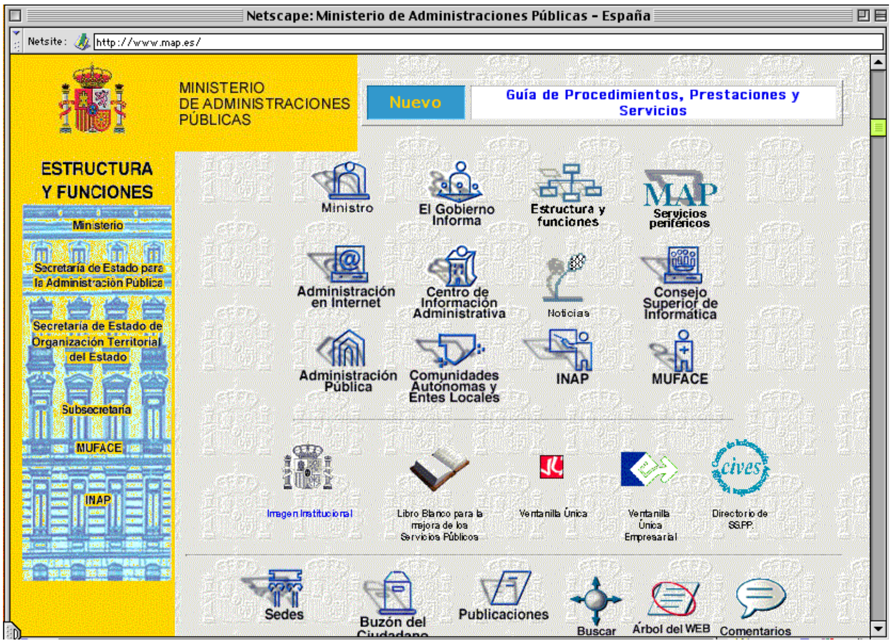
1. Art. 45 of the Law 30/1992, régimen Jurídico de las Administraciones Públicas y del Procedimiento Administrativo Común.

to Contents, page 1

2 the site of High Council for Informatics, the body in charge of the preparation, production, development and application of the Government's policies[2] http://www.map.es/csi/.

Besides information about the most important regulations, their structure and the responsible commissions, it provides useful information about:

- Scope of the administration's areas of activity
- Standards, guidelines and good practice related to products and projects in Information and Communications Technologies. In this field there are important links to products such, for example, the SSD, System of Decision Support for the acquisition of any product or service based on the methodology of discreet evaluation criteria; Metrics Methodologies for applications software development that covers the whole life cycle; the SILICE guidelines for electronic tender bidding; the products and standards derived from the "Unique Window" project (Ventanilla Unica), or the annual Inventory of computer resources of the Public Administrations (REINA and IRIA).
- Special section for the Information society
- Special section about Computer viruses
- News section on Information and Communications Technologies
- Section on current affairs
- Links to the main international Information Technology organisations
- Connections to international organisations in the field of Information Technologies



Site of the Ministry of Public Administration

3 Web site of the Administration: the administrative information hypercentre (AIH) links to most Web sites of the Spanish public administrations, of some of the European institutions and of government institutions from all over the world. Its address is: http://www.map.es/internet/indice.htm.

At the present time the Administrative information hypercentre (HIA) has the appearance of a fundamentally informative website that makes available to all citizens a single access point to the existing Web sites of all the public bodies, such as the General Administration of the State, the Autonomous Administrations of the seventeen Autonomous Communities in which the Spanish territory is divided according to the effective Spanish Constitution of 1975, and of the Entities that form the Local Administration. Those more than 1300 current links are distributed in the following way:

- General State Administration: 140
- Administration of the 17 Autonomous Communities: 160
- Councils, boards and island Town councils: 100
- City councils and other local groups: 800
- Other organisms and institutions of the State: 20
- European union: 35
- Government Web sites of the world: 70

This access service to the information through Internet has links to practically all organisms in the State's General Administration and the Autonomous Communities' Administrations that makes up the Spanish territorial Administration. However, the number of linked City councils is still small, because it amounts to less than 10% of the total of 8,100 Spanish City councils.

Connecting to a single address makes it possible to navigate and access the information pages such as, for example, the State Agency of Tributary Administration (AEAT) for the consultation and the payment through Internet of a number of taxes[3]; the Official Bulletin of the State (BOE), the Official Bulletin of any Autonomous Community or of a province providing this service; the National Institute of Statistics (INE) to find out, for example, the population census of a some municipality; the Museo del Prado to find the program of permanent or temporary exhibitions, any University, or to the page of a community or City council. This service is updated every day as more messages are received from public bodies that request to be included into the HIA system. The information given is

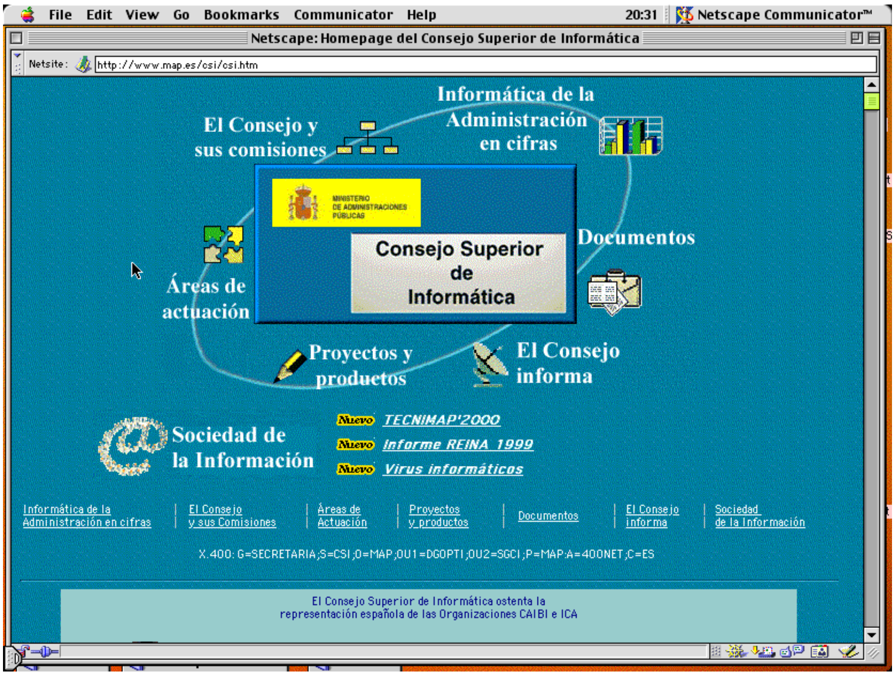
2. The High Council for Computer science was then mandated by the Council of Secretaries to create a centre of attendance and of punctual control of the situation of the centres of the Public Administrations in connection with the Y2K consequences, and in its extraordinary plenary session of May 11, 2000 it has mandated to create a Centre for early virus detection and attacks through Internet for the information, detection and control of the effects that could be derived of them.

3. The second campaign for the tax invoicing (1999) through Internet has already started.

to Contents, page 1

exactly what the information supplier, that is to say, each one of the public organisms, wants to be included, without any editing on the part of the Ministry of Public Administration, only the inclusion, through the creation of a link to the system.

If we look at the last monthly statistic, generated by the product Webtrends, in particular in that of the month of April 2000, we see that the MAP Web sites have received more than 5,100,000 visits; more than 160,000 user sessions, with a daily average of 5,100 sessions; more than 1,000,000 pages have been printed, with an average of about 30,000 pages daily. The users that most visited the pages come from the following countries:



The site of High Council for Informatics

1. Spain 54%
2. United States 28%
3. Mexico 6,25%
4. Argentina
5. UK
6. Colombia
7. Chile
8. Guatemala
9. Peru
10. France 5,2%

The user profile is the following

1. Internal to the public Administrations 50,66%
2. private Companies 41,55%
3. Universities and Education 3,37%
4. Arpanet 1,29%
5. International 2,77%
6. Military 0,36%

The number of accesses oscillates from 194.724 during the weekdays, to 150.484 on Sundays. The page of the HIA, the Administration on the Internet is the most visited one, with 300.000 printed pages monthly.

Current and future work

The current structure of the HIA is hierarchical, based on tree-structures and implies knowledge of the public body in charge of the service that is looked for. In agreement with [Green Book 98] the services that are provided via the Internet are of three types: Information, Communication, interactive/processing. Although still more than 80% of the Internet services are of informative character, gradually the service is becoming interactive. This the one of the objectives of completing the Information Hypercentre and Administrative Processing (HITA).

As for the objective of enriching and improving the product sensibly, and on the basis of the surveys carried out among the users of the Internet services and of the experience of the information services of the Administrative Information Centres, and with the objective of provide the citizens with access to all the Public Administrations without having to worry about knowing on what Ministerial Department a particular organism depends (for example, one that administers, a grant for the opening of a company), or which Public Administration is competent to grant a requested pension. It is a hypercentre for Information and electronic processing, which can be accessed without the need to know the internal structure and line of responsibility of the Public Administration.

A single Internet address (presently this address is a prototype in development and test), will give access to the following services:

- The X.500 Directory of the Public Administration[4]. In this directory all the body of the Spanish Public Administrations are identified, by way of white pages, for the localization of any public body. This will lead to the Directory of Offices of Registration with the postal address, telephone and fax numbers, and e-mail address.
- Contents and topics relating to the electronic information services.
- Organisation chart of the Administration, with a description of functions and services of each administrative organ.
- Complete register of procedures of the participating Administrations and catalogue of application models with standardised forms.

The procedures will be incorporated according to the citizens' interest. The Director-general of the Inspection, Quality and Simplification of the Services takes charge of the classification of the procedures and of the categorization according to

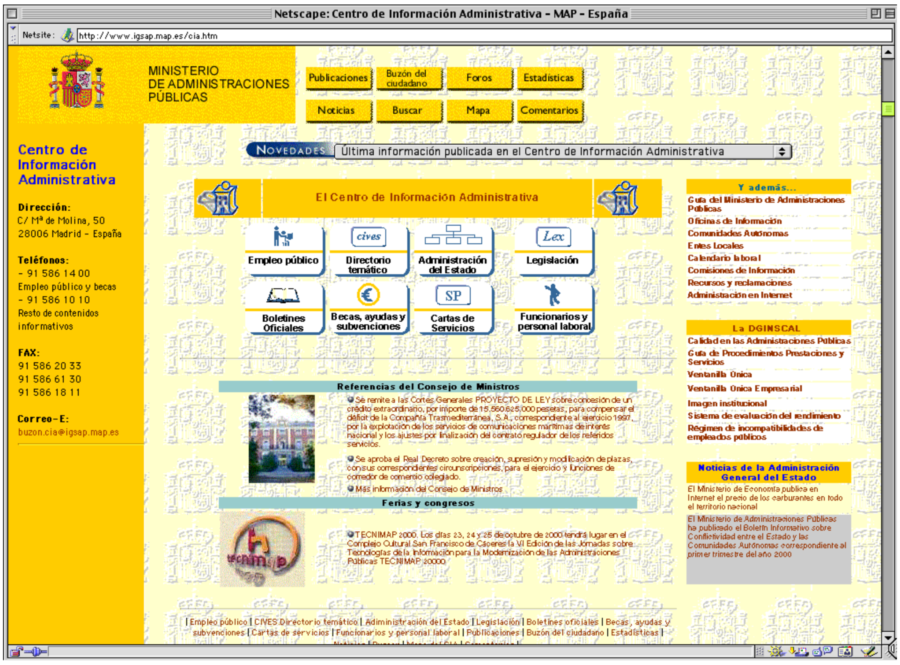
4. Directive that completes the international standard X.500, with unique identifier (UID) for each organism.

to Contents, page 1

its priority, giving more priority to that most requested by the citizens, and leaving for later those for internal use of the Public Administrations. In addition, the following services are gradually being implemented:

- Initiation of electronic processing through a form that can be downloaded from the Web.
- Completion and online-dispatching of the form via Internet.
- Monitoring of the business by means of communications of the concerned administration with the applicant through Internet.

This strategy is being carried out in coordination with the OECD-PUMA initiative for better communication between the Administration and its citizens[5], and a European-wide initiative of the European Union[6], committed to be available by 2003, for all procedures of utility to the citizens of the European Union.



Site of the Administrative Information Centre

In a first phase, in the context of the "PISTA-Ventanilla Unica" project (the unique window)[7], the following standards have been defined http://www.map.es/csi/pg5v20.htm:

- standard for the interface to the Register of entry and exit of the Public Administrations.
- National codification of organisational units (Universal Identifier Descriptor UID).
- X.500 directory of the Public Administrations.
- Records of procedures for electronic processing through the Web.
- Definition of forms of the pilot procedures.

The pilot concluded in September of 1999 and now the bases for the project PISTA-Ventanilla Unica II are being defined. They include important new features[8]. The products are available and waiting for approval by the standards of the High Council for Informatics and their publication as a technical annex in the Official Bulletin of the State. The server for the system resides physically in the Ministry of Public Administration and it is updated locally[9] partly by the participating organisms, and the implementors of the system.

For the success of the installation of this project the commitment of all the participant organisms is necessary, and the highest administrative support so the project extends and finally covers all the public institutions and all the procedures of the Public Administration. Only with strong and determined support of the highest management of the Public Administration will this by seen as a bearing on the improvement of the public services as an improvement by the citizens. A good indication of the implications of that are the commitments published in the "The white book for the improvement of the Public Services. A new Administration for the service to the citizens" [Montes/Yagüe del Valle 99], in relation to the promotion of the Internet services for access to the citizens, and quoted from the third part "The commitments, Chapter 5. An Administration for the Information and Knowledge Society" from which we extract those related to the services provided by the Administration through Internet:

"[...] 3. A multiserve web of the Administration will be created on the Internet that supports voice, data and management, open to the citizens for their relationship with the Administration, and with the other territorial and supranational Administrations. Implementation deadline: June 2004.

4. A single portal of access to the Administration will be created for all the citizens, where all the existing information will

5. The author of this paper represents Spain in the Committee of the OECD-PUMA (Organization for the Trade and the Economic Development – Public Management) for the rapprochement of the relationships between the Administration and the citizens.
6. The author of this paper is also member of the Spanish delegation to the "e-Government" Committee of the European Union, in the initiative "eEurope, a Society of Information for all."
7. Promotion and identification of Advanced Telecommunications Services for "Unique Administrative Window" (Ventanilla Unica administrativa), phase I, a project carried out in collaboration between the Ministry of Development and the Ministry of Public Administrations and co-chaired partly by the Ministry of Public Administrations and partly by the author of this paper.
8. In a very summarized way, the new project seeks, among other, the following achievements: electronic payment; electronic "witness" between different organisms with the objective of avoiding the reiterated petition of the physical documentation that is demanded to the citizens during the procedure; and electronic signature and time stamping.
9. The intervening organisms, by means of a system of public and private keys, have access to their own part of the Web site, may download, and update it.

to Contents, page 1

The Administrative Information Hypercentre

be available, and through which it will be possible to carry out one's administrative business. Implementation deadline: December 2001.

5. A Guide of Services will be developed on the Internet, available to the citizens, where they will find information in real time about the most frequently used administrative procedures and the requirements, documents, procedures, etc. for carrying out in each one of them. Execution term, December 2001."

The public Administration, although presently in development, has as its objective to endow the citizen of an unique, friendly entrance, with a common, simple, intuitive interface by means of an integrated access to the services offered by different public bodies through the Internet. The Portal of the Public Administration intends to become a repository of the knowledge of the different Government departments unifying diverse sources of information and services. The rationale of the installation of the Portal of the Public Administrations lies in the evidence that the new technologies, and in particular the Web, are among the most used means and it will soon be one of the main ways of access for contacts with the public services. The success of the Portal of the Public Services depends greatly on the existence of an efficient search engine (like those in Yahoo, Ole, Ozu, or Altavista,) with indexing capacity for several millions of documents, and the help of an assistant to help users lacking knowledge of the exact location of the services they are looking for.

On the other hand, and as a complement for the more advanced or expert users, the existence of a personalization services agent is recommended that remembers the profile of the Internet users, so as to be able to recommend him, to suggest to him or to suggest to him new products and services, on "one-to-one" contextual web pages, in order to "keep the client." The Portal of the Public Administrations[10] is designed to have attractive pages, with a simple and friendly design with not more than six sections or icons.

The first ones will relate to the Directories of the Public Administration and the products and services for business, organization and for geographical localization. It will consist, therefore, at least, of links to:

- The directory of white pages of the public Administrations.
- All the active URL addresses of the public Administrations.
- The procedures included in the information hypercentre and Administrative Procedure (HITA) and will incorporate those that are added as a result of the PISTA-Ventanilla Unica II project.
- A group of "estrella" ("star") procedures, preferably of horizontal character, obtained as a result of a field study to be carried out in the Administrative Information Centres, and by means of surveys carried out through the Web itself.
- Any other information that is considered of value to the Portal of the Public Administrations is designed to include:
- A suggestions mailbox preformated for automatic distribution to the Organisations addressed by the suggestions.
- All the present contents of the administrative information hypercentre (HIA), operative for the last three years and one of the most visited Web pages of the public Administration, and of the hypercentre data presently in development.
- A news service with special emphasis on the items concerning the Administration the citizens and the possibility of electronic procedures.
- A announcement and advertisement section.

The installation of the public Administrations portal considers the following high-priority aspects as a global project:

- An unified image with a friendly user interface and based on a modern and attractive *design* that serves the double objective of bringing the Administration near to the citizens and "to sell" the products of the Administration.
- The access to the contents with the appropriate *security and quality* measures that observes, on one hand, the privacy, and on the other hand guarantees the access, identification, non repudiation, integrity, availability, reliability and quality of the data.
- The *contents* are based on the coexistence of the information in the public Administrations Web sites, on the results of the work of HITA and of PISTA-Ventanilla Unica II, as well as on the field trials obtained during the first six months of execution of the contract. The high-priority demands of the

10. As with the MAP Web page the redesign is planned to provide it with a more modern and more user-friendly look, with less icons in the first page, and with messages and contents of more impact in the first pages. The remodelling of the page will be done by experts in Web design, experts in knowledge management, and experts in Internet applications.

citizens will be selected from them, and, currently, the services for those that most need them

- A efficient *search engine*, and with capacity of later growth, capable of indexing and to categorizing in the first phase about one thousand five hundred Web sites with a total of several millions of references.

In accordance with the interesting study mentioned in [Hoelscher/Strube 99] it is important to consider the profiles of the users that will carry out the searches, depending on whether they are expert in the material of the searches or are expert in Internet, or if they have a combination of both skills. It is evident that to have a guarantee of high level of success in Internet searches from the Portal of the Public Administrations its design and its search engine must at least, consider the following aspects:

- that the information is included in a controlled way, with semantic analysis, of quality and of relevance of the data.
- whether the user who carries out the searches has previous experience (expert user) in searching traditional document databases.
- If the user has a knowledge on the matters that he is looking for, so that he may apply synonyms and related terms in the search to avoid unsuccessful searches (user expert in the matter)
- If the user has ability in surfing the Internet (junior user).

The need for documental search in the Internet medium still exists and they become more and more important, as the quantity of information offered grows and it is often presented unformatted and without pre-analysis. But the medium has changed, and the modes of access will have to adapt to it and the behaviours of the service providers as well as those of the customers for the service will have to adapt to it. According to [Montes/Yagüe del Valle 99] "As the volume of information stored in the Web increases, it also increases the difficulty of getting back specific information of interest to the user, mainly due to lack of structure and of semantic content in the information stored in that medium." Thus it becomes most important to choose a efficient search engine that is supported by semantic analysis, and the analysis of the citizen's requirements and his social habits are equally important.

The fast and constant adaptation to a medium that evolves in a vertiginous way is no doubt necessary, and the Administration must play a prominent part, not only as referee and producer of standards and best practice as regards the Internet, but also as user of the tool for his internal management; as dispenser of information services, of communication services, all types of communication, and of electronic processing services; as observer of the evolution and the change of the behaviours and social uses; and as catalyst for the creation of employment, the development of the activity of the private sector and the approach to the citizens with the objectives of reducing delays, and supply services of high quality.

Bibliography

[Green Book 98]
Green Book of the Information in the public sector. European Commission. 1998

[Libro 00]
Libro blanco para la mejora de los servicios públicos. Una nueva Administración al servicio de los ciudadanos. Ministerio de Administraciones públicas. 2000. http://www.map.es/libro.htm

[Montes/Yagüe del Valle 99]
J.F. Aldana Montes, M.I. Yagüe del Valle. Consultas sobre la Web. Novática 140, 1999.

[Hoelscher/Strube 99]
Cristoph Hoelscher and Gerhard Strube. Searching on the Web: Two types of expertise. SIGIR'99, Berkeley, CA USA. ACM 1999.

[Savez et al. 98]
Kevin Savetz, Neill Randall and Yves Lepage. MBONE, Multicasting tomorrow's Internet. 1998.
http://www.savetz.com/mbone/

[OECD 00]
OCDE Information Technology Outlook. ICT's, E-Commerce and Information Economy. OCDE, 2000.

[Negroponte 99]
Nicholas Negroponte. El mundo digital. Un futuro que ya ha llegado. Edición especial para TICAP 2000. Ediciones Bailén, Barcelona, 1999.

to Contents, page 1

Challenges in Building up a Client Relationship Channel for Virtual Private Banking

Kornél Szabó

This article provides an introduction to Credit Suisse Private Banking's (CSPB) Internet activities. The first section takes a look at CSPB's Internet strategy before going on to describe the various initiatives under way to implement it. The technologies applied as well as the architectural environment are then outlined. Finally, the article presents the architectural model of the Financial Information Server.

1 E-commerce positioning framework

The increasing interest in Internet-based products and services worldwide is forcing the banking industry to completely remodel its business processes. In response to this development, Credit Suisse Private Banking (CSPB), an independent business unit within Credit Suisse Group, has invested significant effort in creating a new Internet client service channel and in adapting its current business model to meet the requirements of virtual private banking. An e-commerce positioning framework (Fig. 1) has been designed that focuses on the needs of private clients and Independent Asset Managers (IAMs).

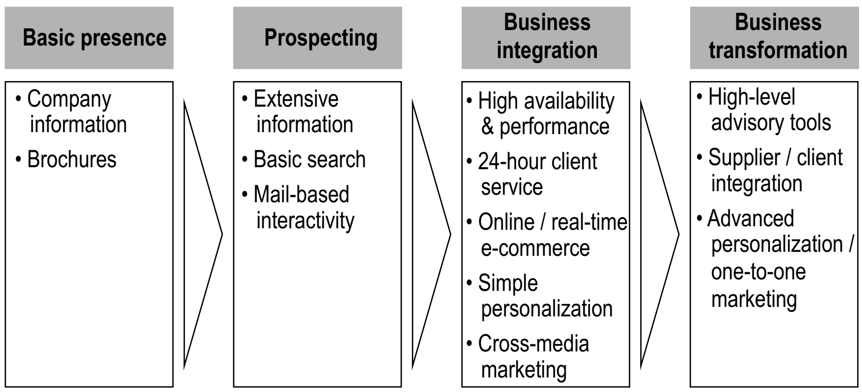


Fig. 1: CSPB's e-commerce positioning framework

The major point of entry for all CSPB Internet initiatives is the Web site and community portal www.cspb.com, which CSPB aims to develop into an integrated financial service channel for existing and potential clients. The user should be accompanied from information browsing through to final transaction processing and the point of sale.

2 CSPB Internet initiatives

2.1 Initiative I – Internet presence and provision of information

When CSPB established its web presence in 1997, its focus was on providing information on the company itself, financial products and services, corporate news in the form of press releases, as well as on highlighting events sponsored by CSPB. It also gave the user the opportunity to contact a CSPB representative via the Internet using a digital contact sheet. In 1998 CSPB started to establish the Internet as a new client relationship channel. Its Internet content architecture (Fig. 2) was based on three zones: the public zone, the member zone, Investors' Circle, and the private zone, Online Banking.

In order to broaden and deepen its web presence, CSPB has initiated the following:

- cross-media marketing: systematic communication of www.cspb.com as a brand in every marketing campaign (brochures, posters, newspaper ads, TV ads, etc.)
- Internet marketing campaigns (banner ads, Web site submissions, doorway pages, response control)
- country-specific Web sites with integrated access to current local banking structure and country-specific application interfaces

2.2 Initiative II – interactive investment tools and alliances

Several interactive investment tools have been implemented to speed up the business transformation process. They support users in their investment decisions and provide a transparent view of financial, insurance and real estate products. Different third-party data providers (e.g. insurers, banks, real estate agents, market data providers) work closely with CSPB to enhance the tools with up-to-date and comprehensive product

Dr. Kornél Szabó received his PhD at the Multimedia Laboratory of the University of Zurich in the field of virtual reality based information systems. Since mid-1998 he has been working for Credit Suisse Private Banking in the Virtual Banking department and is in charge of web-mastering and IT project management.
kornel.szabo@cspb.com

to Contents, page 1

Zone	Public	Member (Investors' Circle)	Private (Online Banking)
Access	• virtually unlimited • restricted according to country selected (-> cookie checking)	• via user name and password • permanent access for CSPB clients; expiry after 30 days for non-CSPB clients	• via contract no., password and SecurID or strike list
Content	• market quotes and news • interactive investment tools • products & services • jobs and career opportunities • events sponsored by CSPB • media releases • company presentation • link to other CS business units and CSPB country websites	• research reports • technical research • equity and bond recommen-dations • personalised portfolio tracker	• DIRECT NET (account information, payment services, stock prices, placement of stock exchange orders) • youtrade (online trading, stock prices, research information)

Fig. 2: CSPB's Internet content architecture

information. Before they could create tools that people really use, CSPB had to factor issues such as usability and interfacing to applications and data feeds into the application design.

2.3 Initiative III – community building and personalisation

A further step towards providing a client-focused solution and building up a community on the web was the establishment of a sophisticated area within the CSPB portal called Investors' Circle (IC). The client is able to log in to a password-protected area offering information such as in-house research reports, equity and bond recommendations provided by CSPB analysts as well as a portfolio tracker. A dedicated login server controls access to the IC. CSPB clients can request permanent access, whereas non-CSPB clients can obtain access for a trial period of 30 days. The portfolio tracker is the first personalised application within the IC that allows users to configure application functionality and data views in line with their individual needs.

2.4 Initiative IV – consistency, optimization and stability

Various initiatives have been launched in line with the generally accepted usability guidelines for web applications such as speed, consistency and simplicity. Web site design guidelines have been elaborated and strictly applied to all new web applications with the aim of presenting a consistent look and feel on all CSPB Web sites. Web pages have been optimized in size and performance. Special tools to run heavyweight load tests have been integrated into the development process in order to make the applications more stable. In addition, a set of internal as well as external tool-monitoring systems have been set up to increase system up-time. The establishment of company-wide technology and system set-up standards has been very important with regard to the influence of human factors on infrastructure stability and security. The use of business-wide application and infrastructure configuration guidelines have helped to shorten the development process and time to market.

3 Technologies in operation

3.1 System architecture and components

CSPB's Internet infrastructure is operated mainly by the Credit Suisse business unit. Human as well as system and networking resources are shared across the Credit Suisse Group. This allows people to fall back on experience gained in former projects and enables them to rapidly pass on their expertise internally. CSPB endeavours to carry out technological gatekeeping and to adopt new technologies in its Internet-related projects at an early stage. Sun and UNIX constitute the computing platform, while the Netscape Enterprise web server and Oracle database system form the core Internet server infrastructure. Applications are written in Java, Perl, C/C++, SQL and served to the web client as HTML/Javascript or Java Applets. The various Internet web servers are hosted within a demilitarized zone (DMZ) protected by carefully configured firewalls, proxies and routers. New Internet applications have to pass a security audit before becoming operational.

3.2 Design and management of data feeds

Almost all CSPB Internet applications involve data feeds provided by internal or external data sources. In order to overcome the various data feed requirements, CSPB decided to design and implement a core architecture and software framework for data feed processing. The development was carried out in-house and driven by the following requirements:

- to gain flexibility in developing financial information services
- to offer a standard API for client applications
- to enable cross-referencing of various in-house as well as external data or information
- to provide a rich set of search and look-up functions
- to implement a modular and scalable system architecture to quickly fulfil growing demand for existing and new services
- to make it easier to integrate additional data sources
- to build up internal know-how of data feed architectures and applications

The goal is to offer financial information in-house to various CSPB Internet applications from a central server infrastructure called Financial Information Server (FIS). A conceptual view of its basic functional building blocks is shown in Fig. 3.

The core technological components used to implement the FIS are Java, Java Beans, SQL/Oracle, XML and CORBA.

to Contents, page 1

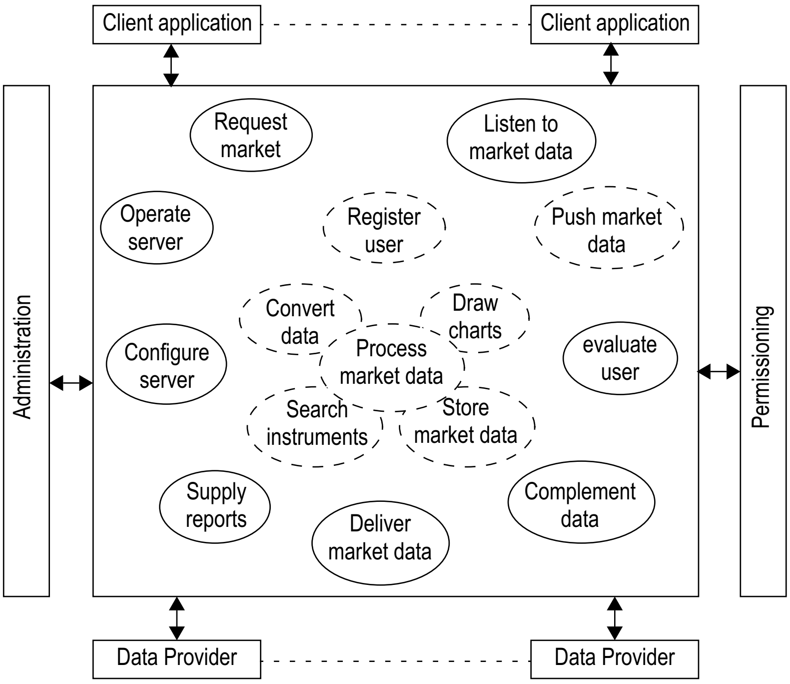


Fig. 3: Basic functional building blocks of CSPB´s Financial Information Server

4 Lessons learned and critical success factors

Two of the key factors in rapidly implementing the Internet strategy and becoming a portal for virtual private banking are high-level management support and close cooperation between business and technical staff members. Short iteration cycles and a common cultural ground help to overcome many difficulties and to implement the targeted company-wide business transformation process. Thinking in terms of workflows, interfacing and usability is essential for the establishment of a sustainable web presence.

Early consideration of operating issues during the system development process, early integration of project stakeholders, consistent change and version management and the use of technical as well as operational coordinators enable the smooth configuration and integration of new applications into the existing infrastructure.

In order to build up and operate a coherent application environment, the design and implementation of an interfacing framework is important. Consequently, all applications have to respect this framework and offer adequate support for functional and content interfacing. Security issues and scarce human resources may be handled more efficiently by relying on corporate technology standards and by following the KISS (Keep It Simple and Stupid) principle in application design. For Internet-based private banking applications where the user primarily does not interact with a physically present relationship manager, usability engineering issues have to be an integral part in application and page layout design. The various code and layout optimization steps described briefly in point 2.4 helped to increase interaction performance significantly. HTML code could be optimized in order to achieve 30–50% (partially even 100%) average speed-up in download time and performance. Users provided a very positive feedback to these improvements in the Web site.

5 Conclusion

Credit Suisse Private Banking has recognised the importance of the Internet as a new, promising client relationship channel. It has consequently initiated the appropriate business transformation processes to meet the new, challenging business and technological requirements. A new virtual banking department has been created with the task of driving the Internet strategy and providing CSPB with a high degree of visibility on the web.

Various initiatives have been launched and are still ongoing aimed at building up and establishing an Internet portal for virtual private banking. On the technology side, usability, interfacing and data feed issues are regarded as critical success factors. On the business side, management support, short communication paths and the elaboration of workflow standards play a key role in achieving rapid success and a strong market focus.

to Contents, page 1

Designing Electronic Network Organisations for ICT-Enabled Health Care Networks

Ryan R. Peterson, Martin Smits and Ronald Spanjers

This paper describes the design and development of electronic network organisations in health care. The strategic drivers, design and ICT infrastructure of electronic health care network are outlined. Based on an in-depth investigation of electronic health care networks, this paper summarises the main lessons learned and the critical success factors. The implications for research and directions for health care practice are discussed.

Keywords: Electronic Network Organisation, Health Care, Information and Communication Technology Infrastructure, Stakeholder Partnerships, Network Designs and Development Stages, Case Studies.

1 Introduction

The network economy is challenging traditional well-established health care institutions to develop new patient-oriented models and invest in information and communication technologies. Once a cottage industry[1] of physicians, hospitals, medical centres and consultants, health care is now experiencing the value of integrated services and the collaborative advantage of networking. Economic, social, political and technological forces have driven physicians, consultants, hospitals and medical centres to develop electronic networks relationships, and 'e-care' is rapidly becoming a norm of quality [Peterson/De Wit 99].

While much is presumed and predicted about electronic network organisations in Health Care, little empirical evidence exists as to their drivers, design and development, and even less directions exist for health care practitioners. Based on an in-depth investigation of electronic health care networks, this paper addresses these issues and provides several guidelines.

2 Framing Electronic Network Organisations

A network organisation is distinguished from a classical organisation by the intensity, density, multiplexity, and reciprocity of inter- and intra-organisational ties, and a shared value system defining stakeholder[2] roles, responsibilities and relationships. An electronic network organisation is characterised by non-hierarchical, long-term commitments; multiple distributed stakeholder roles and responsibilities; independent, yet interdependent decision-making; and an ICT-based network infrastructure [Ribbers/Smits 99].

1. Cottage industry: a business or manufacturing activity carried on in people's homes. *(ed.)*
2. Stakeholder: a person with an interest or concern in something. *(ed.)*

If electronic network organisations are going to proliferate and become the dominant organisation type on the emerging economic landscape, they must exhibit unique features that are particularly well adapted to the new environmental exigencies. Situational factors are described by the strategic drivers and enabling conditions, and the main network dimensions refer to the design and processes of the network organisation and ICT infrastructure (Figure 1).

The main dimensions of electronic network organisations are [Peterson et al. 00]:

- *Strategic drivers:* The network objectives and motives of organisations and stakeholders involved in the network.
- *Enabling conditions:* The conditions that enabled or stimulated the emergence and formation of the network.
- *Network Design:* The structuring of the network responsibilities, decision-making units, and coordination mechanisms

Ryan Peterson is assistant professor in Information Management at the Tilburg University. His research interest include Information Governance (Ph.D.), IT Strategy and Organization, Electronic Business and Network Organizations, Multimedia Technology and Innovation Management. He started his research career in 1996 after receiving his M.Sc. in Organizational Technology and Management by the University of Twente. <r.r.peterson@kub.nl>

Ronald W. L. Spanjers started in 1992 as a Financial Consultant in the Bosch Medicentre (general hospital) for the Social and Financial Department. From 1998 he has been working as a Manager in the University Medical Centre Utrecht for the Division Obstetrics, Neonatology and Gynaecology. In the time left he is working on his Phd research at Tilburg University in which organisational and IT development in hospitals are main topics. <r.w.l.spanjers@azu.nl>

Martin Smits is associate professor of Information Systems at Tilburg University. His research focuses on strategic information systems planning, group decision making, inter-organisational systems and new organisational forms. He was involved in the development of computer based teaching and management information systems in health care. He holds a Ph.D. in Medicine by the University of Groningen. <m.t.smits@kub.nl>

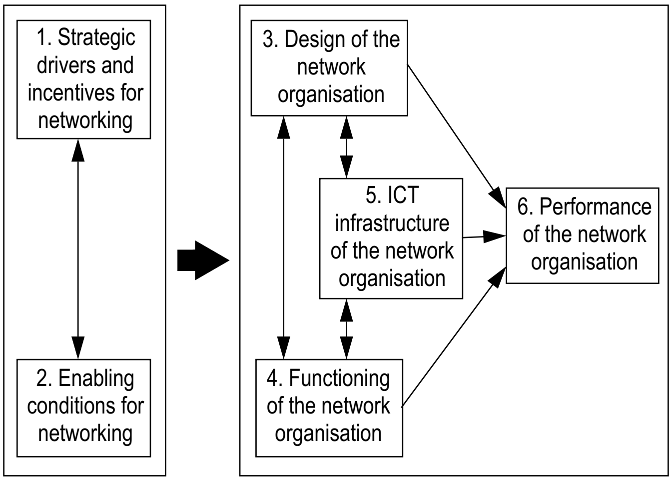


Fig. 1: Framing Electronic Network Organisations

- *Network Processes:* Network business activities and transactions.
- *Network ICT Infrastructure:* The reach, range, and standardisation of the ICT infrastructure.
- *Network Performance:* Organisational and ICT impacts and effects of the network.

3 Electronic Health Care Networks in Practice

In this paper we analyse three cases of electronic health care networks. The first two cases, the Roessingh Rheuma Network (RRN) and the Bosch Medicentre Network (BMN), are located in the Netherlands. The third case study, the Renal Telemedicine Network (RTN), is located in Australia. Summaries of the cases are presented in Table 1.

Bosch Medicentre Network

On the 6th of January 1990 the Willem-Alexander Ziekenhuis and the Groot Ziekengasthuis merged into the Bosch Medicentrum. The Bosch Medicentrum is a general hospital with a capacity of 780 clinical beds, 1,900 full time equivalents, 2,600 employees and 140 medical specialists. In seeking improved efficiency and effectiveness, the Bosch Medicentrum in 1996 started a reorganisation evolving from a facility management into a product-line management organisation structure.

Roessingh Rheuma Network

Roessingh Research and Development is a research unit of the Roessingh Concern and employs approximately 40 people. The Roessingh Concern has approximately 140 beds and approximately 40,000 rehabilitation treatments per year. It is one of the largest rehabilitation centres in the Netherlands. The Rheuma network was formed when a proposal was submitted to the Commission for Chronically Ill Patients to formalise and institutionalise communication lines between Medical Spectrum Twente (MST) and local clinics, and Leiden University Medical Centre (LUMC) and local clinics. This proposal was submitted in August 1998 and was a joint effort of both MST and LUMC.

Renal Telemedicine Network

The Queen Elizabeth Hospital (TQEH) provides a comprehensive range of specialist and diagnostic treatment services to the immediate community in western metropolitan Adelaide as well as country areas. TQEH's Renal Telemedicine Network (RTN) commenced in June 1994. Over 75% of patients are supported on haemodialysis and in South Australia the majority of these are located in "satellite" centres. Information and communication technology applications and infrastructure were installed at its four renal dialysis centres at TQEH Woodville and Wayville (10 km from Woodville) in September 1994, and at North Adelaide (8 km) and Port Augusta (300 km) in February 1995. RTN dialyses a total of 145 patients at these four centres, with each patient normally dialysing three times per week and attending an outpatients clinic once every two months. The network organisation also cares for 29 patients who dialyse at home.

4 Electronic Networks in Development

In general, the case studies provide ample evidence that electronic health care network organisations are in a constant flux, driven and enabled by both external opportunities and internal needs. The electronic health care networks come in different forms and shapes, and have different functions.

Electronic health care network organisations develop through different phases of maturity and networkability as they migrate and grow synergistically (Figure 2). However, while ICT enables the formation of professional networks in health care, it is ultimately the health care network partnerships that determine the direction and development of the network. Internal and external partnership building are especially important to the transformation of electronic network organisations in health care.

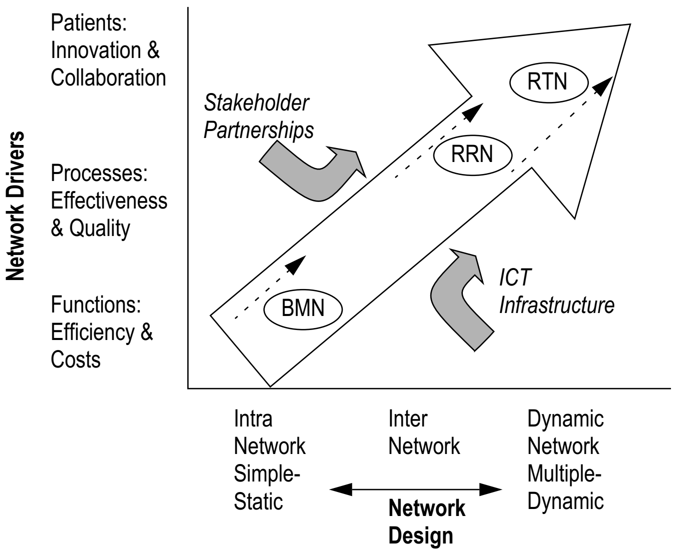


Fig. 2: Electronic Health Care Networks in Development.

to Contents, page 1

Network Dimension	Bosch Medicentre Network	Roessingh Rheuma Network	Renal Telemedicine Network
Strategic drivers and incentives for networking	Reduce costs, improve efficiency, no loss of quality in hospital health care.	Improve efficiency and effectiveness of rheumatology services in order to meet patients' needs. Formalise communication and develop shared expertise.	Improve efficiency and effectiveness of renal dialysis services in order to meet patients' needs. Improve communication and educate specialists.
Enabling conditions for networking	*The evolution from a facility management into a product-line management organisation structure.* Efficient and effective control of hospital organisations depends on ability to determine the relation between input and output.	*The demand and supply mechanisms regarding rheumatology knowledge across the network.* The demand and supply mechanisms regarding ICT knowledge across the network. Partnership building	*The demand and supply mechanisms regarding renal dialysis services across the network.* The demand and supply mechanisms regarding telemedicine applications across the network. Management of change. Partnership building
Design of the network	*Intra-Network.* Hospital-personnel (2,600) and medical Specialists (140) work together to attend to the patients needs. Hospital middle management, functional operators and medical specialists and automation coordinator.	*Inter-Network.* Separate responsibilities for rheumatology services and ICT services. Different functional roles and levels: "sponsor", "network coordinator", "participants/ users". Health care participants	*Dynamic Network.* Separate responsibilities for renal dialysis services and telemedicine technology services. Different functional roles and levels: "sponsor", "network coordinator", "participants/ users", "technology integrator". Health Care and ICT participants
Functioning of the network	*Internal medical information support.* Technical support, implementation and central computing facilities are outsourced.	*Network and stakeholder management.* Provision of telerheumatology services across the network. Leveraging of rheumatology expertise across the network. Demand and supply of multimedia network technology.	*Network and stakeholder management.* Provision of renal dialyses services across the network. Leveraging of renal dialyses expertise across the network. Demand and supply of telemedicine technology.
Network ICT infrastructure	*Dependent view of ICT infrastructure.* HISCOM information systems are used. Low reach and range of infrastructure.	*Enabling view of ICT infrastructure.* Network standardisation. Moderate reach and range of infrastructure.	*Enabling view of ICT infrastructure.* Network standardisation and multi-tier architecture. High reach and range of infrastructure.
Performance of the network	*Reduced costs without loss of quality in hospital health care.* Networking is reaching beyond organisational boundaries in transmural care projects.	*Improvement of inter-institutional collaboration and communication.* ICT flexibility and reliability. Efficiency and effectiveness improvement of rheumatology services. Stakeholder satisfaction. Redefinition of stakeholder roles.	*Improvement of inter-institutional collaboration and communication.* ICT flexibility and reliability. Efficiency and effectiveness improvement of renal dialysis services. Stakeholder satisfaction. Institutionalisation and growth.

Table 1: Network dimensions and key findings.

Both the technical and the social infrastructure are important for the design and growth of electronic network organisations in health care. Moreover, network management, focusing on stakeholder partnerships and social coordination, is pivotal to the successful development of electronic health care networks [Peterson/De Wit 99], [Smits/Van der Pijl 99].

Based on our investigation, the following lessons and critical success factors are drawn on the drivers, design and performance of electronic health care networks:

- *Strategic Drivers*
  Strategic drivers for electronic network organisations are improvement of efficiency and cost-effectiveness, process quality and effectiveness, and primary care process for meeting patients' needs and patient-information-streams. However, these strategic drivers differ over time as electronic network organisations transform (see Figure 2). Different "rationalities" exist for improving the efficiency and effectiveness of health care services. From an internal perspective, inter-organisational collaboration and expertise development are emphasised. In the external environment, the will to meet patients' needs is underscored.

- *Enabling Conditions*
  The important lesson learned in all cases is the critical role played by management and the process of managing stakeholders' needs and expectations, and inter-organisational change. Managing the demand and supply of care and technology is a key enabler of network organisations in health care. On one hand, there is the need to share and collaboratively develop health care expertise. On the other, there is also the need to apply ICT to facilitate the efficient and effective delivery of health care services.
  Proper attention to organisational, political and human issues can not be overstated enough in the development of successful network organisations, especially in health care where professionals carry the "power to innovate". While ICT may provide the conditions for networking, it is the organisation, its professionals and management that ultimately drive networking. Attention for stakeholder motives and expectations is therefore critical.

- *Network Design*
  Regarding the design of network organisations in health care, the case studies cover a spectrum from inter-organisational (the RRN and RTN cases) to intra-organisational and (the BMCN case) network organisation (see Figure 2). Furthermore, each case covers a different phase of growth: piloting (the BMCN case), learning (the RRN case) and growing (the RTN case).
  Interesting is also the growth of ICT infrastructure reach and range during these different phases. This growth in ICT is readily recognised in the RRN and RTN cases in which telemedicine technologies are applied in more organisation functions as the network grows. During these phases, ICT applications are modified and redesigned to meet the needs of users and the different clinical, educational and administrative services. Flexible ICT infrastructures are therefore

required in developing dynamic electronic network organisations, driven by the need to innovate and collaborate to meet patient demands.

- *Network Processes and Functioning*
  In the functioning of network organisations in health care different processes are distinguished. Health care transactions, in the form of patient-information streams and clinical communication, between health care service providers are at the core of network processes. Network management and coordination are likewise important and organised through both formal and informal mechanisms in which key stakeholders take part.
  Important network coordination mechanisms include stakeholder committees, joint decision-making, informal and formal communication and stakeholder involvement. ICT-mediated coordination in the form of electronic databases and video conferencing are especially used in Inter Networks and Dynamic Networks.

- *Information and Communication Technology*
  With regard to the role and impact of ICT in electronic network organisations in health care, the case studies indicate that ICT plays an important role in each of the network dimensions as described above. However, it is the network constituency that needs to recognise, adopt, implement and exploit the potential opportunities provided by ICT. An "enabling view" of ICT is associated with a higher level of networkability. While ICT enables the formation of professional networks in health care, it is ultimately the health care network that drives and determines the acquisition and application of ICT in the network.
  The cases indicate that ICT requires constant modification to meet the specific needs of the health care network functions, and that the ICT supplier to the network organisation needs to be actively involved in the different stages of network formation and professionalization. The flexibility of ICT, expressed as the infrastructure reach and range, is important to the development of electronic network organisations. Equally important are the multi-tier architecture standards that need to be agreed upon by the different stakeholders.

- *Network Performance*
  The ability to describe and measure the performance of the network increases when a higher phase of growth has been reached. In the piloting phase, performance measures are described in general terms, repeating the mission statement. In the learning phase, performance is likewise described in general terms, with a focus on stakeholder expectations and networking agreements. Stakeholder roles are redefined and the performance is assessed in terms of stakeholder satisfaction. In the growing phase, stakeholder roles have been institutionalised and the "benefits" of networking become clear. These include health care efficiency and effectiveness gains, professionalization and expertise development, and stakeholder satisfaction.
  The financial performance of the network remains difficult because the relation between input and output in a health care organisation is hard to determine. Moreover, "traditional" cost-benefit analyses of network organisations in health care are sub-optimal because they fail to account for all the (inter-/intra-organisational) changes that occur as a result of networking.

5 Future Directions

As described in the foregoing sections, this investigation provides a number of lessons learned for designing electronic network organisations in health care. Electronic network organisations are in a constant flux, driven and enabled by external opportunities and internal needs, addressing both organisational and ICT infrastructures. As the network economy evolves and electronic network organisations transform, research is on-going in this field, and currently other cases in the health care sector are being studied in the second phase of a research programme on network organisations (see also http://nefeti.kub.nl).

As part of a long-term research programme on network organisations in health care, and other industries, our research endeavours are geared at:

- analysing and understanding emerging business models of network organisations, and the supporting and shaping role of ICT infrastructures and applications;
- providing directions and guidelines for developing and migrating towards network organisations, and the implications and requirements for ICT infrastructures.

In particular, future research is focused on identifying and explaining electronic network designs of high performance network organisations in different phases of development. Only then can we truly begin to understand the design and dynamics of electronic network organisations in a global network economy, and provide design tools for developing successful electronic network organisations.

References

[Peterson/De Wit 99]
R. R. Peterson and D. De Wit: ICT-Induced Innovation in Health Care: A Challenge to the Management of Multimedia Technology. *Proceedings of the 7th European Conference on Information Systems 1999*, Copenhagen, Denmark.

[Peterson et al. 00]
R. R. Peterson, M. T. Smits and R. W. L. Spanjers: Exploring ICT-Enabled Network Organisations in Health Care: Emerging Practices and Phases of Development. *Proceedings of the 8th European Conference on Information Systems 2000*, Vienna, Austria.

[Ribbers/Smits 99]
P. M. A. Ribbers and M. T. Smits: Network Enterprises of the Future and Enabling ICT Infrastructures. Internal Working Paper NEFETI 99.01, 1999, Department of Information Management, Tilburg University.

[Smits/Van der Pijl 99]
M. T. Smits and G. J. Van der Pijl: Developments in Hospital Management and Information Systems. *Proceedings of the 32nd Hawaii International Conference on System Sciences 1999*, Hawaii.

Strategies and IT Integration in Outpatient Doctors' Networks

Stefan G. Gfrörer, Markus Raupp and Franz Schober

We analyse doctors' networks particularly with a view on firm level network strategies and the supporting information technology integration. Our main conclusions are that the viability of doctors' networks critically depends on trust building mechanisms like the restriction of the network in size and complexity and the application of fair profit allocation rules. Concerning information technology the implementation and use of highly integrated inter-organizational systems appears most promising. We propose a system architecture that integrates information technology along the medical, the business and the communication systems dimension.

Keywords: Outpatient health care, cooperating physicians, doctors' network, integrated interorganizational information system.

1 Introduction

Because of its already very high and still increasing costs the German health care system has been under heavy debate during recent years. As about 60 percent of the overall health care costs are caused by outpatient health care, new and more efficient organizational solutions of outpatient health services have been searched for. One proposal that plays a major role in the current debate is the establishment of doctors' networks. In the meantime some pilot networks have been established in Germany [Kassenärztliche Bundesvereinigung 99], but it is too early to draw firm conclusions from these first experiences.

The German outpatient health care system is embedded into a strong regulatory framework. Important parties are the medical association ("Kassenärztliche Vereinigung"), and the statutory health insurance companies which are admitted to the system. The medical association has mainly two objectives: firstly to represent the interests of the member physicians in the society, and secondly to act as a clearing organization between physicians and insurance companies.

With more and more financial constraints on the statutory scheme, the system exhibits some serious deficiencies. As the total financial volume for redistribution to the physicians is fixed, the system invites for opportunistic behaviour on the doctors' side. This typically implicates a hidden, but nevertheless fierce competition between the physicians to increase the volume of services provided and in consequence jeopardizes the cost efficiency of the overall system [Milde 92].

Therefore, other solutions have been proposed. One of the proposals centres around the idea of doctors' networks, where several physicians from complementing disciplines establish a relatively stable and long-term cooperation. A core element of a doctors' network, hereby, is the joint treatment of a patient by several legally independent network members for a lump fee which has to be allocated to the treating physicians by some network-internal mechanism. In this sense a doctors' network is quite distinct from and goes far beyond other forms of cooperation between physicians [Schober/Gfrörer]. The lump fees are negotiated between the doctors' network and the medical association or even directly with some or all of the statutory insurance companies on a case-specific as well as a network-specific level, which is possible in Germany since 1st January 2000.

While the doctors' network approach would reduce the problem of opportunistic behaviour in the current system, it poses new problems. One is the potential rivalry between the doctors within a given network that would transport the problem of opportunistic behaviour only to another level. A second question concerns the competition between networks and the creation of competitive advantages of one network against rival networks. Thirdly, and most importantly, the network approach must be accepted by the patient, i.e. must be more attractive than the current situation with individually operating physicians. Besides many other factors, the solution to the three problems is also impacted by the proper use of information technology in a doctors' network. As we will argue, a common and highly integrated inter-organizational information system (IOS) constitutes an important strategic resource that helps to

Stefan G. Gfrörer was research assistant at the Department of Information Systems, University of Freiburg, Germany, and is now with Itellium System & Services GmbH, Frankfurt/M, Germany. Subject of his research is the support of doctors' networks by IT and the resulting network efficiency and stability. <gfroerer@vwl.uni-freiburg.de>

Markus Raupp is research assistant at the Department of Information Systems, University of Freiburg, Germany. His line of research centres in the field of strategic planning in network organizations and its implications for the design of inter-organizational information systems. <raupp@vwl.uni-freiburg.de>

Franz Schober is professor and director of the Department of Information Systems, University of Freiburg, Germany. His current research interests are concerned with information technology and organizational design and with knowledge management. <fschober@vwl.uni-freiburg.de>

add value to the services rendered to the patient, to keep cost under control and to position the network against rival networks.

The aim of our paper is to evaluate the key factors that are critical for the success of doctors' networks both from a strategic and information technology (IT) perspective. In chapter 2 we apply a strategic framework, that we have developed for the analysis of different types of networks in general [Raupp/Schober 00], to evaluate firm-level strategies of physicians participating in doctors' networks. Chapter 3 deals with the role of IT more specifically, and proposes an overall architecture for an IOS in doctors' networks. Chapter 4 concludes with a summary and open questions for further analysis.

2 Strategies in Doctors' Networks

2.1 General conditions for strategy definition in networks

The viability and success of any network arrangement essentially depends on the individual contributions to the overall objectives of the network, respectively on the investments of each network member. Bargaining power exploitation by some participants in order to proactively influence the individually appropriable profit shares can negatively impact the investment incentives of the other network members and in consequence can compromise the overall efficiency of the network arrangement [Schober 99]. It is therefore important to classify different network arrangements contingent to their inherent bargaining power properties. Following [Raupp/Schober 00] networks and their inherent bargaining power distribution can be characterized by network size and network topology. These dimensions essentially determine the necessary degree of formalization in the underlying network coordination structure, which itself is primarily determined by means of explicit or even implicit contractual specifications. Joint investment plans for IOS as well as plans for an inter-organizational process integration constitute important parameters of the network coordination structure [Raupp/Schober 00].

Network size describes the number of members of a network. Network topology is determined by the distribution of economic linkages between the network members. In symmetric networks most members interact with each other, either directly or via a joint activity such as a professional association. A typical example for a symmetric network arrangement is the medical association, which coordinates some common interests for its members. In asymmetric arrangements, there exist some network members which have significantly more bilateral linkages than others, like in buyer-supplier networks where some buyers do business with several suppliers, and where little interaction occurs between the suppliers.

By means of cooperative game theory it can be shown, that in symmetric network arrangements bargaining power imbalances decrease with increasing network size. In contrast, in asymmetric arrangements bargaining power imbalances intensify with increasing network size [Raupp/Schober 00].

Comparable to buyer-supplier networks also doctors' networks may exhibit substantial asymmetries in bargaining power, particularly if some network participants take a gatekeeper's role as it is designated in the German proposal [Gfrörer et al. 00]. Asymmetric arrangements like doctors' networks are much more prone to opportunistic exploitation of bargaining power, trust exposure, and the corresponding negative impacts on investment propensities and the total network efficiency.

In the following we will focus our analysis on doctors' networks where gatekeepers take the role of network coordinators and represent the network as a whole against the external environment (e.g. patients, the medical association, or the insurance companies). Typical tasks of gatekeepers are the negotiation of case-specific or network-specific lump fees with the medical association or directly with the insurance companies as well as the routing of patients through the network. Gatekeepers are usually represented by highly reputable general practitioners. Obviously, because of the resulting information asymmetries these actors exhibit a dominant bargaining position in doctors' networks. Nevertheless, the participating physicians are not direct competitors but possess complementary skills and aim to cooperate on a stable and long-term basis. Furthermore, these arrangements include common investment strategies, e.g. into sophisticated medical technologies, but also into proprietary and highly integrated IOS to gain sustainable competitive advantages against rival networks. Because of these characteristics it is almost impossible to institutionalize a network coordination structure on basis of *ex ante*[1] specified contracts and rules. In consequence of the incompleteness of contracts and the asymmetric distribution of bargaining power, doctors' networks require trust-building mechanisms to ensure the long-term viability of the arrangement [Gfrörer et al. 00]. These mechanisms are strongly related with the firm-level network strategies and will be discussed in more detail in the following chapter.

2.2 Firm-level network strategies

Firm-level network strategies, i.e. in our specific context the strategies of each individual physician participating in a network arrangement, are an integral part of the overall business strategy and concentrate on the positioning of each member *vis-a-vis* the other network members. It is important, that the firm-level strategies fit with the objectives on the overall network-level respectively with the coordination structure of the underlying type of network arrangement.

Firm-level network strategies comprise two dimensions: profit sharing strategies and resource sharing strategies [Raupp/Schober 00]. Both dimensions can adopt either a competitive or a cooperative connotation. Profit sharing strategies are concerned with the allocation of revenues, investments, and costs and depend on the strategically motivated application of bargaining power. Competitive profit sharing strategies are characterized by a full exploitation of bargaining power while cooperative profit sharing behaviour comprises limitations of bargaining power exploitation, for example, by means of an *ex ante* implementation of fair profit allocation rules including fair sharing of joint investments [Schober 99] or an *ex ante* restric-

1. *ex ante*: based on forecasts rather than actual results.

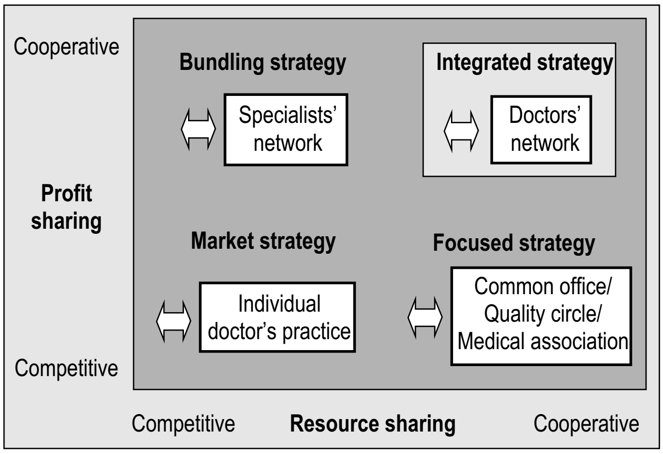


Fig. 1: Typology of firm-level network strategies.

tion of the network size. The resource sharing strategy focuses on the management of resource positions in inter-organizational arrangements. Cooperative resource sharing strategies focus on the creation of joint network specific competencies whereas competitive resource strategies strictly emphasize the protection of core competencies for each network member.

Based on the cross comparison along the cooperative respectively competitive orientation of the two strategic dimensions four distinctive categories of firm-level network strategies can be distinguished: integrated strategy, focused strategy, bundling strategy, and market strategy. They correspond with different types of possible network arrangements, see Fig. 1.

Physicians participating in a doctors' network can benefit from a mix of cooperative resource sharing strategies and cooperative profit sharing strategies [Gfrörer et al. 00]. Pooling of medical core competencies and patient information hand in hand with joint investment into medical technology and integrated IOS constitute central objectives of a doctors' network. Chapter 3 will discuss the aspect of IOS integration in doctors' networks in more detail. Furthermore, the creation and nurturing of social and managerial network competencies including the ability to develop trust, to share risk and to enhance collective learning are of central importance. Typically, a key strategic asset in doctors' networks stems from the high degree of inter-organizational process integration. The productivity of an individual doctor in the network and the investment incentives of the other network members are reciprocally dependent. As a consequence, cooperative profit allocation strategies to strengthen mutual trust and investment incentives are critical for the long-term viability of doctors' networks. Trust building requires a renunciation of bargaining power exploitation by dominant network members, particularly on the side of the gatekeeper. The restriction of overlapping competencies in doctors' networks hand in hand with the *ex ante* limitation of the network size constitute important trust-building mechanisms. If asymmetric bargaining power is preserved, the weaker partners tend to underinvest and the resulting economic position is worse for all actors in the network [Schober 99].

In contrast, physicians participating in a "common office", a "quality circle" or in the "medical association" are typically characterized by "focused strategies". Apart from fundamental differences between these types of networks, in all arrangements the participating physicians primarily concentrate on the cooperation in a very specific domain of interest. In all other domains the physicians remain competitors, i.e. operate independently with individually owned patients and individual service provisions. Common office arrangements share mainly physical resources, whereas quality circles and the medical association focus on the sharing of information resources [Schober/Gfrörer]. In the latter case the cooperative resource sharing particularly focuses on data aggregation and fee settlement. But because of the competitive nature of all three types of network arrangements, the resource sharing is typically restricted to distinct domains of inter-organizational collaboration. Furthermore, the inherent competitive elements induce low incentives for joint investments into integrated IOS. Therefore, in these arrangements voluntary constraints on bargaining power exploitation typically are not required and consequently also trust-building mechanisms play a less important role.

"Bundling strategies" are especially suited for networks cooperating on a temporary basis. They apply to physicians who do not cooperate within the network on a permanent base, particularly to highly specialized physicians whose services are purchased by a doctors' network on demand (specialists' networks). Since there are no incentives for resource sharing in this case, we also would not assume joint investments into highly integrated IOS.

In our specific context, the "market strategy" and the related organizational arrangement "individual doctor's practice" do not reflect a network arrangement but serve as reference points in our framework.

With the emphasis on trust-building mechanisms, on a high degree of IOS integration and on the creation of a network-specific knowledge base, doctors' networks stand in sharp contrast to the other forms of cooperative doctors' arrangements [Gfrörer et al. 00], [Schober/Gfrörer].

3 IT Integration in Doctors' Networks

The knowledge about the patient and his or her treatment constitutes an essential competitive asset of a physician. This knowledge resides not only in explicit patient records, but also in implicit or "tacit" knowledge about each individual diagnostic and therapeutic case. Tacit knowledge cannot be codified and therefore also not transferred to other physicians except if they work very closely together. Here we see an immense source of enduring competitive advantage for doctors' networks. Sharing of knowledge and collective learning are key characteristics in doctors' networks. Because of its limited transferability, this knowledge base constitutes a barrier for patients to move to other networks or physicians outside the network, even if explicit medical records would be transferred.

Because of the tacit component, knowledge in a doctors' network cannot be represented by electronic patient records and IT-based applications as such. Yet, IT can significantly enhance the knowledge creation process even on the tacit level [Schober 00]. In addition, IT can also help to improve the economic efficiency of a doctors' network.

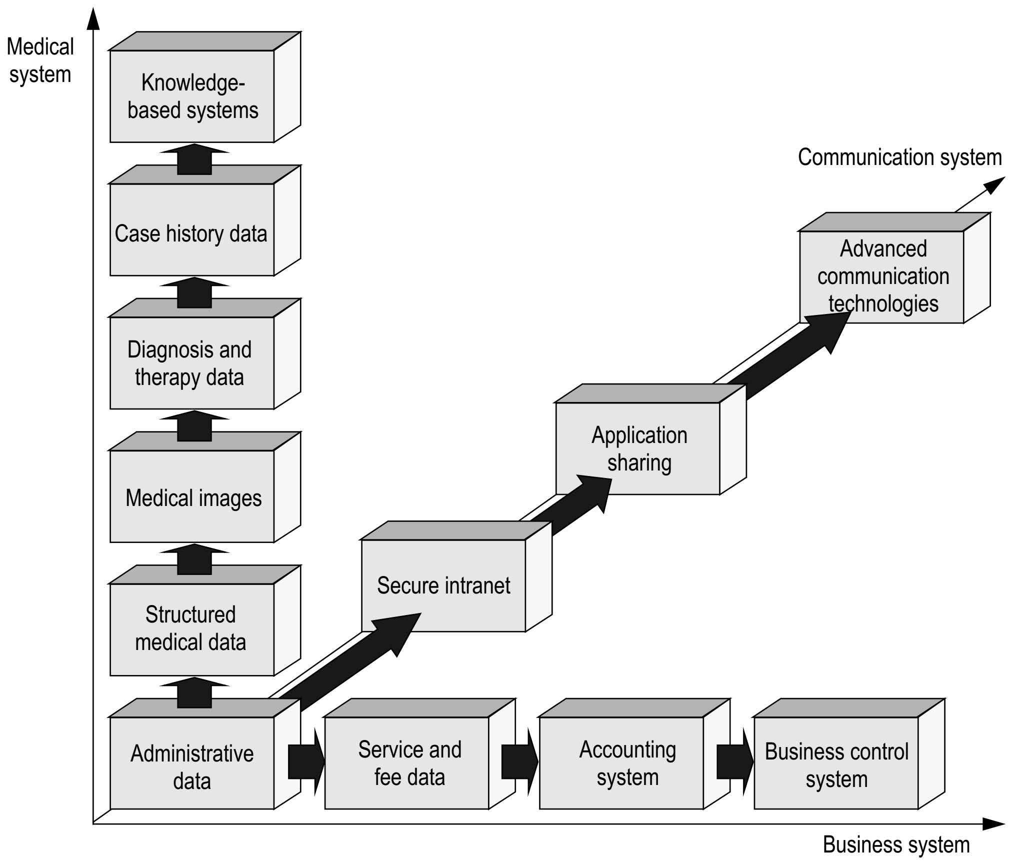


Fig. 2: Architecture of an integrated IOS.

In the following we propose an architecture for an integrated and IT-based inter-organizational system (IOS) for doctors' networks which supports both the knowledge creation process as well as the search of economic efficiency. Fig. 2 gives an overview of the IOS architecture. It is characterized by step-wise system integration along three dimensions: integration of the medical system, integration of the business system and integration of the communication system.

The core of each dimension of integration is a shared database with administrative data, e.g. general patient data or general data of the participating physicians including specific competencies and installed medical technologies. This administrative database is shared by all other applications along the three dimensions of integration.

Along the dimension of integration for the medical system we successively add structured medical data including laboratory measurements, images (e.g. roentgen or tomographic images), data on diagnosis and therapy and on patient case history. The highest integration level of the medical system refers to knowledge-based systems which support the physician in diagnosis and therapy. Although the intensively discussed medical expert systems have not found wide appreciation in practice there exists a promising application potential for intelligent checklists or for the use of case-based reasoning in comparing an actual diagnostic with previous cases of other patients. However, unstructured elements in the electronic patient record (e.g. dictated or written reports) raise several problems in data communication, data management and data administration. Therefore, metadata technologies that address the management of large bodies of text and facilitate the discovery and interconnection of similar medical cases play an important role [Panko et al. 99]. The extraction of case history data represents a task far beyond the capability of administrative database systems. This implies to build up an extensive and formalized case history which certainly would constitute a major competitive advantage of a doctors' network.

The integration along the dimension of the business system starts with the compilation of network-specific lists for available services and fees based on elementary settlement runs. A more sophisticated common accounting system represents a next step of integration. It would contain all important business data like investments into medical, office or information technology, salaries, material costs, and purchases from services outside the network. The accounting module should also include an accepted algorithm for redistribution of fees between different service providers in the network and the calculation of the overall remuneration for each physician. The accounting system will ultimately serve as the basis for a common business control system which provides and compares various indicators to measure the economic efficiency of the network. The business control system requires also a tight link to the medical system, particularly to deliver specific indicators like cost by disease or treatment categories.

A key element of the integration along the dimension of the communication system is system security, particularly because of the sensitivity of patient data. System security comprises the control of access rights (which are basically at the patient's side and have to be verified by the patient) and the secure encryption of data, preferably already in the database but in any case when passing through public networks. Application sharing allows several physicians to share the same application. This is particularly important for the update and retrieval of patient data, but also for most of the other applications in the IOS, e.g. certain knowledge-based systems for diagnosis support. Architecture-independent Java applets coupled with mark-up languages (SGML and its subset XML) represent cost-efficient options to work simultaneously at different locations with one set of data and applications [Panko et al. 99]. Advanced communication technologies such as video-conferencing improve the real-time communication between the doctors in the network. The use of e-mail and the World-Wide Web may also serve as a platform for communicating directly with the patients, hospitals and other third parties.

The IOS components in Fig. 2 are certainly far from being complete and should only give some important directions of integration. Nevertheless, a high degree of IOS integration along the medical, business, and communication dimensions surely is a necessary prerequisite for knowledge creation in tightly coupled doctors' networks, but by itself it is not a sufficient condition. Social and managerial network competencies including capabilities for conflict resolution and trust building represent further flanking mechanisms for efficient knowledge creation in doctors' networks.

4 Summary and Conclusion

Within this paper we have analysed two critical dimensions for the success and the long-term viability of doctor's networks: cooperative resource and profit sharing behaviour on the firm-level on the one hand and a high degree of process and IOS integration on the other hand. The latter aspect essentially requires shared goals between the participating physicians that are embedded in a trust-based overall network coordination structure.

We have identified doctors' networks as arrangements that are exposed to bargaining power asymmetries and therefore require the institutionalization of trust-building mechanisms. These include the restriction of the size and complexity of the network as well as fair rules for fee, investment, and cost allocation. The restriction in size contradicts with the configuration of the current pilot networks in Germany which seem to be quite large [Kassenärztliche Bundesvereinigung 99]. Furthermore, the pilot networks seem to involve overlapping competencies of the participating physicians [Kassenärztliche Bundesvereinigung 99], a central inhibitor of trust-building in network arrangements.

Concerning inter-organizational information technology we have voted for several reasons for a highly integrated solution. One reason stems from the long-term nature of the cooperation and the importance of trust. The second reason relates to the sharing of common resources as one important objective of doctors' networks. Integrated IOS make this sharing more efficient. Another reason is the creation of a knowledge base. A fourth reason, finally, concerns the patient. Only a highly integrated IOS provides the necessary integrity in the sense that it presents the network to the patient as one entity rather than a collection of more or less independent physicians.

Our analysis leaves a series of questions open. One is the coexistence of individual physicians and physicians organized in networks. Another open question concerns the cooperation between doctors' networks, hospitals, and highly specialized physicians. Also cooperative arrangements between doctors' networks and insurance companies constitute a topic for further investigation. Some legal aspects were touched only very briefly, particularly the sensitive question of patient data protection. Therefore, much research in the area of network organizations in health care is still needed.

References

[Gfrörer et al. 00]
S. G. Gfrörer, M. Raupp and F. Schober: "Restructuring the German Outpatient Health Care System: An Economic and IT Perspective," in *Proceedings of the 8th European Conference on Information Systems* (ECIS 2000), H. R. Hansen, M. Bichler and H. Mahrer, Eds. Vienna, July 2000, pp. 1261–1268.

[Kassenärztliche Bundesvereinigung 99]
Kassenärztliche Bundesvereinigung, Ed., "Projekte zur Weiterentwicklung der ambulanten Versorgung im Überblick," in *KBV Kontext* 9, Köln, 1999.

[Milde 92]
P. C. Milde: *Institutionenökonomische Analyse alternativer Krankenversicherungssysteme – Das Beispiel der Gesetzlichen Krankenversicherung und der "Health Maintenance Organization"*, Diss. Hamburg, 1992.

[Panko et al. 99]
W. Panko, J. Silverstein and T. Lincoln: "Technologies for Extracting Full Value from the Electronic Patient Record," in *Proceedings of the 32nd Hawaii International Conference on System Sciences* (HICSS-32), S. F. Nunamaker and R. H. Sprague, Eds. Los Alamitos, CA: CD-ROM distributed by IEEE Computer Society Press, January 1999, 9 pages.

[Raupp/Schober 00]
M. Raupp and F. Schober: "Why Buyer-Supplier Chains Differ: A Strategic Framework for Electronic Network Organizations," in *Proceedings of the 33rd Hawaii International Conference on System Sciences* (HICSS-33), S. F. Nunamaker and R. H. Sprague, Eds. Los Alamitos, CA: CD-ROM distributed by IEEE Computer Society Press, January 2000, 10 pages.

[Schober 99]
F. Schober: "Kostenallokation für interorganisationale Informationssysteme," in *Electronic Business Engineering / 4. Internationale Tagung Wirtschaftsinformatik* 1999, A. W. Scheer and M. Nüttgens, Eds. Heidelberg: Physica, 1999, pp. 135–146.

[Schober 00]
F. Schober: "*Organisationales Wissen und Informationstechnik," in Freiheit und wettbewerbliche Ordnung – Gedenkband zur Erinnerung an Walter Eucken*, B. Külp and V. Vanberg, Eds. Freiburg: Haufe, 2000, pp. 501–521.

[Schober/Gfrörer]
F. Schober S. G. Gfrörer: "Network Organizations in Outpatient Health Care: Conditions for Viability and Success," Forthcoming in *Proceedings of the Joint Nagoya/Freiburg Seminar on "Aging Society"* (in Japanese).

Continuing Medical Education and Professional Development

Marcial García Rojo

This article reviews the goals of medical education, that must be redefined broadly due to the presence of new scenarios, where new approaches, such as the implementation of logical reasoning process, have to be contemplated. We describe how this can be achieved through the use of Internet and other authoring tools. Evidence-based medicine is essential when physicians are dealing with literature. Frequency of use of Internet and the quality of the information are analysed. In a second part, we review the advantages and disadvantages of virtual congresses, and the impact of new technology terminology in medical literature. In a near future, the use of medical integrated workstations will be common; they will allow fast and efficient information searches, and they will be essential to access patient previous history, related cases, or finding the best choice of treatment.

Keywords: Continuing Medical Education, Medical Informatics, Internet, Evidence-Based Medicine, Quality of Information.

1 Introduction

Continuing medical education (CME) is part of the process of lifelong learning that all physicians undertake. It has traditionally been viewed by the medical profession in terms of updating their knowledge. Nowadays, in order to practise effectively in the modern National Health System, doctors need skills that extend their medical knowledge beyond updating. Thus, hospital doctors and general practitioners have now accepted responsibility for both continuing medical education and professional development (CPD).

Similarly to other professional environments, medical knowledge, necessary for teaching or education purposes, is usually stored in computerized system. This process has been possible with the advent of electronic versions of specialized journals, congress proceedings on CD-ROM, the development of authoring tools with multimedia support (e.g. anatomy atlas, or diagnosis support expert systems). Certainly, the most important extension of medical informatics, since it appeared at the end of 70s, has occurred with the use of the Internet, with its useful hypertext text links in web pages, that allow easy development and dissemination of information integrating text, pictures, video, sounds, etc.

Thus, the rapid popularity of the World-Wide Web has changed the ability of medical doctors and patients to access a vast amount of information. At the same time, there is also a fast advance in basic and clinical sciences, diagnostics and therapeutics techniques, which professionals need to integrate into their daily medical practice.

Marcial García Rojo is a M.D. Ph.D. He also holds a Master in Computer Science. Pathologist at the Hospital of Ciudad Real (Spain). Founding member of the Informatics Special Interest Group of the Spanish Society of Pathology. Member of the Board of Directors of Spanish Health Informatics Society (SEIS).

1960s: Research Systems (ECG, diagnosis).
1970s: Management, department, image (TC Scan), Mycin.
1980s: Reports, outpatients, Clinical I.S. and databases, artificial intelligence.
1990s: Integration, communication, vocabulary, evaluation

Table 1: History of Main Efforts in Health Informatics

As shown in table 1, that summarizes the historical evolution of applications in health informatics, main efforts were initially directed, with limited success, towards management process or experimental expert systems instead of towards solutions for education of health professionals.

This evolution did not correspond with the real needs of medicine professionals, who are usually closer to current Internet solutions developed by universities, pharmaceutical companies and communication companies, amongst others, that offer free access to updated information, including tools for search of medical articles (available or not in Medline), experts consultation, and so on.

[Collen 94] summarizes in the following points the reasons for the delay in the use of digital computers in the practice of clinical medicine :

1. There was a 10 years delay in information technology in health systems
2. It is a complex environment where simple solutions, like those in bank companies, are not applicable
3. The need of technology able to manage complex information (QMR or Dxplain)
4. Inertia, the fear of pre-established rules

The increasing interest in new information technologies amongst medical professionals, in parallel to the extensive use of the Internet, can be corroborated by Health on the Net (http:/

• National Library Medicine, Medline	http://nml.nih.gov/
• British Medical Journal (BMJ)	http://www.bmj.com
• Medscape	http://www.medscape.com/
• Intelihealth	http://www.intelihealth.com/
• Mayo Health System	http://www.intelihealth.com/
• Mayo Health System	http://www.mayohealth.org/
• Health On the Net	http://www.hon.ch/
• The Lancet	http://www.thelancet.com/

Table 2: Most frequently requested web sites by doctors in 1999.

/www.hon.ch) surveys, that show how, in Europe, in 1999, 46% of doctors were Internet users, and, although 95% of them evaluate as useful the information available in Internet, a 7% of expert users consider that this information should improve greatly.

These surveys also show the most popular web sites in the medical community, that are shown in table 2. It is obvious that sites offering biomedical information (electronic versions of traditional medical journals and Internet servers containing information about diseases) were the most requested in 1999.

The noticeable increase of papers related to information technology, together with the increasing presence of medical journals in Internet, has significantly expanded the interest in this area amongst health professionals [LaPorte et al. 95].

Using multimedia techniques in Internet, the University of Sydney has achieved a 100% use of these technologies by medical students, comparing with a 25% of students using them at the beginning of 90s [Kidd et al. 93].

This paper focuses on the different solutions of continuing education directed to current practitioners who generally have not received any specific training in information techniques. Although this technology is being simplified, most authors agree to include a postgraduate programme in medical informatics in order to obtain an efficient use [Coiera 98].

Every effort in continuing medical education should be preceded by the evaluation of the cognitive process (reasoning and decision making), instead of creating a simple data collection with the classical textbook format [Patel et al. 00].

[Coiera 98] proposes the following Decalogue of essential clinical informatics skills:

1. *Understand the dynamic and uncertain nature of medical knowledge, and be able to keep personal knowledge and skills up-to-date*
2. *Know how to search for and assess knowledge according to the statistical basis of scientific evidence*
3. *Understand some of the logical and statistical models of the diagnostic process*
4. *Interpret uncertain clinical data and deal with trial and error*
5. *Structure and analyse clinical decisions in terms of risks and benefits*
6. *Apply and adapt clinical knowledge to the individual circumstances of patients*
7. *Access, assess, select and apply treatment guidelines, adapt them to local circumstances, and communicate and record variations in the treatment plan and outcome*
8. *Structure and record clinical data in a way appropriate for immediate clinical tasks, for communication with colleagues, or for epidemiological purposes*
9. *Select and operate the most appropriate communication method for a given task (e.g., face-to-face conversation, telephone, e-mail, video, voice-mail, letter)*
10. *Structure and communicate messages in a manner most suited to the recipient, task and chosen communication medium.*

In 1996, the percentage of Spanish doctors using computers was below 40%, a figure similar to those observed in France or Portugal, but considerably lower than the 80% observed in British or German practitioners [Ceusters/Prorec-Be 98].

2 New goals in education

[Ludvigsson 99] has proposed the following guidelines to redefine the goals of medical education:

- *Understanding biomedical concepts related to disease mechanisms The Flexnerian or "reductionist" model of teaching, where the body is considered as a machine with organs that can be repaired by specialists in that specific system, should be avoided.*
- *Developing interpersonal and hands-on skills, including forming productive partnerships with patients and health care team members, and demonstrating appropriate professional values.*
- *Applying a logical reasoning process to solve individual or community problems and to critically review new information.*
- *Accessing information resources appropriately to support high quality practice.*

New technologies allow students to proceed at their own pace and create flexible learning environments. Advanced web technology allows programmed access to learning modules and tests. "Virtual" tutorials are conducted by threaded discussion groups. These methods hold great promise for allowing lifelong learning through accessing up to date electronic knowledge resources on the web and CD ROMs [Neame et al. 99].

In addition, during their clinical years, students can use web based technologies for keeping a learning log to record their studies and to accumulate a personal log of cases or problems seen, procedures watched or undertaken, and skills acquired for certification purposes a system often described as portfolio based learning [Snadden 99].

3 Authoring tools

In this section we have grouped together those solutions implemented with conventional programming languages (C++, VisualBasic, Delphi, etc.) that usually need an installation process in each computer. In medical informatics there has been a noticeable predominance of Mumps and Open M languages.

to Contents, page 1

GROUP	
Clinical history:	61
Information systems	16
Dentistry	10
Radiology (Reports and Imaging)	8
Medical office	6
Clinical laboratory	4
Anatomic pathology	4
Cardiovascular	3
Other specialities	10
Management	13
Education	4
Anatomic or diseases atlases	2
Medical literature	2

Table 3: Medicine software available in 1997

These tools are used for education programs in specific areas, protocol guided systems, or computerised physician decision support.

For instance, StrokeNet was developed to simplify emergency attention for strokes, through a recompilation, on one hand, of present symptoms and signs, and on the other hand, by generating diagnostic and therapeutical recommendations based on the most recent research studies [Moehr 00].

In the last few years these tools have been integrated into the so-called groupware applications, such as Lotus Notes. These allow a direct link between the student and the professor, as well as numerous utilities, like discussion forums with the simultaneous interaction of multiple students.

Therefore, this kind of solution will probably be integrated in health networks and information systems in primary care and hospital services, as additional and supplementary components of Internet technology. An example of this integration can be observed in the projects included in HealthNet, in Canada, that since 1996, has developed medical informatics projects, including collaboration programs between different participating groups.

[Keravnou et al. 97] have grouped intelligent medical systems into:

- ***Protocols and guidelines:*** *protocols for medical procedures and therapies, clinical guidelines, health care processes;*
- ***Automatic diagnosing and decision support tools:*** *knowledge acquisition and learning, decision support theories, diagnostic problem solving, probabilistic models and fuzzy logic;*
- ***Temporal Reasoning and Planning:*** *planning and optimising of therapies, patients management, global health care planning, planning environments;*
- ***Natural language and terminology:*** *medical dictionaries, automatic abstracting, information retrieval, communication, multilingual dictionaries, lexicons;*
- ***Image and signal processing:*** *image interpretation, pattern recognition and identification;*

We have collected a list of 78 commercial professional health programs, distributed throughout Spain in 1997. Only 4 programs (5%) were related to continuing medical education (table 3).Campbell and Johnson conducted a Medline search, collecting all available abstracts of articles addressing multimedia computer aided learning (n=258) and multiprofessional learning (n=92) written in 1985–98. They conclude that most publications (63%) in Medline that represent multimedia computer aided learning are project descriptions or position statements, rather than reports of research (34%) including qualitative or quantitative methods used to systematically investigate the topics. Furthermore, references to established educational principles were infrequent (24%) and few abstracts (7%) referred to educational theory. Research is dominated by quantitative methods of questionable validity and utility. Relevance to practice is centred on teaching, but with minimal consideration of established educational principles or theories [Campbell/ Johnson 99].

4 Evidence-Based Medicine

Evidence-Based Medicine (EBM) is supported by three main pillars: a) the proliferation of clinical research projects related to new technologies, above all, drugs; b) the development of clinical research methods; and c) the important increase in clinical documentation. In summary, it deals with all medical literature [Marimón 99].

Therefore, the main goal in EBM is the conscious, explicit and judicious use of the best evidence available to support clinical decisions in patient care. In other words, considering relevant clinical investigations, including precision and accuracy in diagnostic tests, prognostic markers influence, and efficiency and security in therapeutic, rehabilitating, and preventive measures [Bravo Toledo 00].

Available evidence search tools are, on one hand, secondary (filtered) sources, such as briefing journals with critical reviews (e.g. Evidence-Based Medicine journal, edited by the American College of Physicians, with publications in Internal Medicine, Primary Care, Paediatrics, Surgery, Psychiatry, Obstetrics and Gynaecology), the Journal Clubs (e.g. ACP Journal Cub), CAT bank (Critical Appraisal Topics http://cebm.jr2.ox.ac.uk/docs/catbank.html), and resources to find the former (such as Bandolera, also available in Spanish at http://www.infodoctor.org/bandolera/); on the other hand, specialized database such as Cochrane Library.

Although evidence-based medicine requires new skills from physicians, including searching literature, new technologies use should be evaluated prudently. [Verhoeven et al. 00] performed a randomised comparative study to determine which literature retrieving method is most effective for general practitioners (GPs): the printed Index Medicus; Medline through Grateful Med; or Medline on CD-ROM. They observed that in the period 1994–1997, the printed Index Medicus was the most effective literature retrieval method for GPs. For inexperienced GPs, there is a need for training in electronic literature retrieval methods [Verhoeven et al. 00].

to Contents, page 1

5 Internet

The Internet has become an essential resource in continuing medical education. The main services we can find on the Internet, suitable for continuing education, are virtual congresses, specialized courses, clinical cases, virtual patients, discussion forms, or literature revisions, among others.

Most of these services are free and available to general public, if not already dedicated to a restricted professional group. They can be published to inform other people about the institution's activities or to record scientific meetings, even those not related to the Internet.

Frequency of use

An interesting study was done in Norway, at the end of 1998, analysing physicians' continuing medical education and their information-seeking behaviour, including their use of the Internet. 72% of Norwegian doctors had Internet access and 24% had access either at work, at home or both. Doctors with access both at work and at home used the Internet significantly more often and found it of greater professional value than did the other groups. A smaller proportion of general practitioners, compared to other groups in the profession, had access to the Internet. About half (48%) of Norwegian doctors use the Internet in a professional context. Research-oriented, male doctors, 30–49 years old, indicated the highest activity on the net. Authors conclude that, for the time being, it appears that the net widens the gap between doctors who actively seek new professional knowledge and those who do not [Nylenna 99].

Quality of Information in Internet

[Sandvik 99] evaluated the Internet as a source of information about urinary incontinence. 75 web sites providing interactive information about incontinence were analysed, along with 25 web doctors, and two news groups. Popularity indexes were measured using Hotbot and Altavista indices, according to number of links to web sites. Excellent information about urinary incontinence was found on the Internet, but the number of links to a site did not reflect the contents quality. Patients may get valuable advice and comfort from using interactive services. Well known organisations (societies, foundations, and journals) will probably offer better information than universities, hospitals, and clinics (labelled "professionals") and "commercial" sites.

In English-speaking countries, professional colleges (like royal colleges in U.K.) are responsible for providing a framework for continuing professional development; setting educational standards; and monitoring, facilitating, and evaluating activities for their members [du Boulay 00]. In Spain, this area is not well structured, and it is usually not well coordinated by

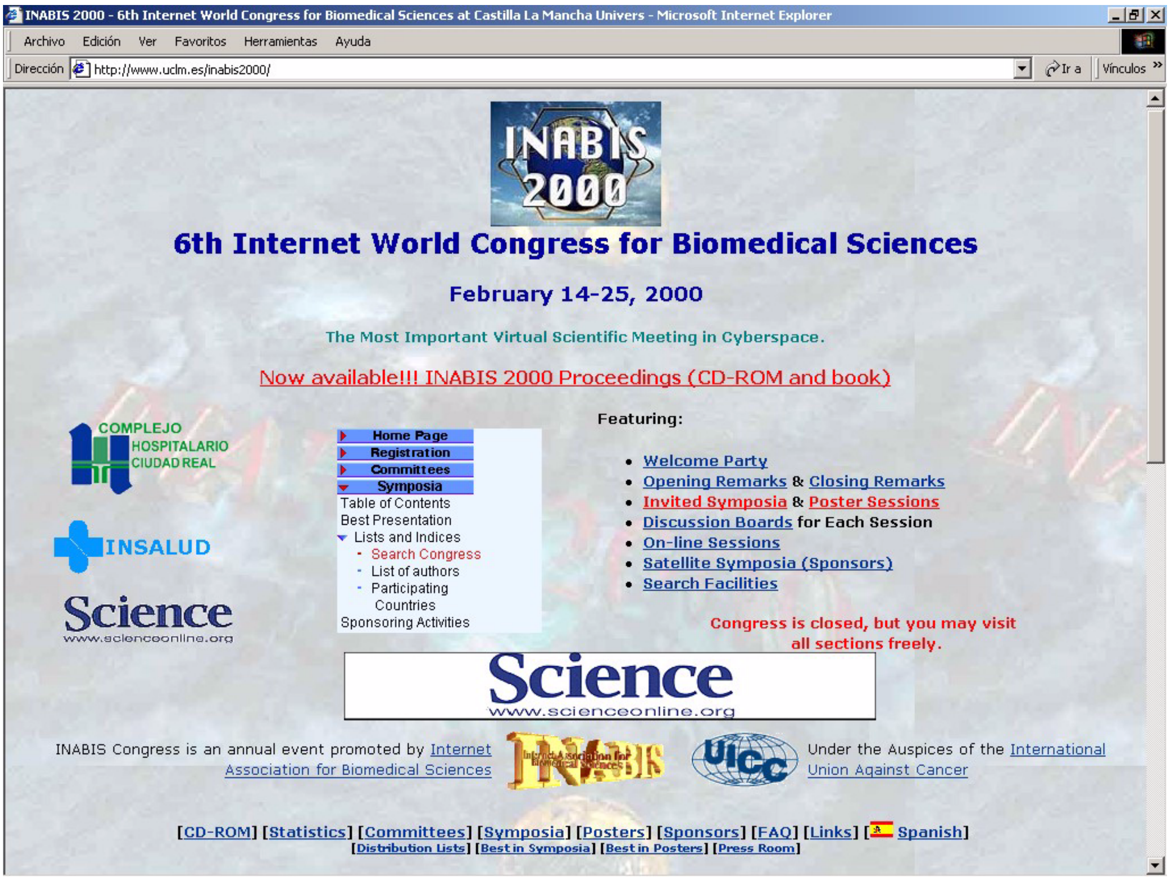


Fig. 1: World on-line congress INABIS 2000

to Contents, page 1

They present larger opportunities to meet new colleagues with common interests
The ability to promote professional relationships even when congress is over
The ability to publish worldwide information about the city or institution where event is organized
The lower costs of organization
The avoidance of registration and accommodation costs
The automated evaluation of abstracts (via web or e-mail)
The easy coordination of efforts of multiple organising societies
Presentations can be directly modified by the authors
The unlimited number of scientific papers
The unlimited number of registered delegates or visitors
The information is available for an unlimited period of time
Efficient search facilities available
Active phase can be as long as needed

Table 4: Advantages of Virtual Congresses

.

Scarce interest of commercial drug companies in sponsorships
Slow connections to the Internet
Little active participation in discussion forums
Specialized Internet programmers needed
Some scientific journals does not admit presentations previously published on the Internet

Table 5: Disadvantages of Virtual Congresses

several professional bodies, such as regional medical colleges, scientific societies, and the Health Ministry.

Cryveillance.com estimated that there are more than 2,100 million web pages [Anonymous 00]. Although no specific data are available about health web sites in Spain, Compuware company, using WebCheck utility, has analysed the main companies present in the Internet, and found that 65% of Spanish web sites have important defects in design and access speed. The most common deficiencies were missing attributes, pages under construction, and broken links.

Physicians usually access general search engines in their searches: Alta Vista, Hotbot, Northern Light, Infoseek, Excite, Galaxy, Yahoo, or MSN. They also use specialised servers for Medicine, such as Achoo, Cliniweb, Health AtoZ, Healthfinder, Karolinska Institute, Medical Matrix, Medical World Search, Medsite, and OMNI. Some of the most visited sites are: MedNets (http://www.mednets.com/index.html), Martindale's Health Science Guide (http://www-sci.lib.uci.edu/HSG/HSGuide.html), MedWeb, de la Emory University (http://www.medweb.emory.edu/MedWeb/), MedWeb Plus (http://www.medwebplus.com/), MedMark (http://medmark.org/) and BioMednet (http://www.bmn.com/).

The most frequently used search engines in Spain, as expressed by the Media General Study, are Yahoo (21,6%), Terra (28,9%) and Altavista (17,1%), followed by, Ozú (5,9%), Lycos (5,1%) and Excite (2,2%) [Anonymous 00]. Some frequently visited specialized Spanish web sites are Diario Médico (http://www.diariomedico.com), Recol Network (http://www.recol.es/), MedicinaTV.com (http://www.medicinatv.com), and Saludalia for Physicians (http://www.saludaliamedica.com/).

Year	CME Search	MeSH Search
1991	369	54
1992	357	43
1993	385	41
1994	405	77
1995	462	91
1996	474	108
1997	541	118
1998	539	120
1999	613	132
2000 (-oct)	347	70

Table 6: Number of articles in free searches about Continuing Medical Education (CME) terms, compared to searches performed using MeSH terms Medical Informatics and Continuing Education

Virtual Congresses

We started the organization of virtual congresses on Pathology in 1996, (http://www.conganat.org/), and recently, our group was responsible for the organization of the VI World Congress on Biomedical Sciences in Internet – INABIS 2000 (http://www.uclm.es/inabis2000/), see Figure 1. This experience allowed us to evaluate the growing interest that health professionals have in this new type of meetings. In tables 4 and 5 we have summarised the main advantages and disadvantages found in the organization of virtual congresses

6 New information and communication technologies impact in medical literature

We have recently studied the use of terms related to information and communication technologies, using the on-line database [PubMed 00] to perform Medline searches in scientific journals.

In our first approach, we analysed the use of Medline Browser MeSH terms (http://www.ncbi.nlm.nih.gov/entrez/meshbrowser.cgi) such as "Education, Continuing" and "Medical Informatics"). We obtained a total of 1,201 articles, compared with a total of 11,376 articles when non-specific searches about Continuing Medical Education (CME) were performed.

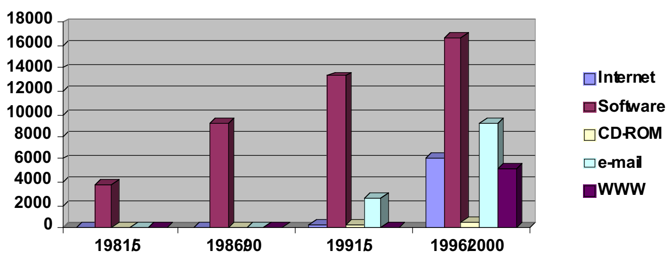


Fig. 2: Frequency of technical terms in scientific articles between 1981 and 2000.

to Contents, page 1

Main Topic	Number of articles		
	1998	1999	2000
Total	91	112	61
New technologies in education	21	24	12
Internet	19	37	22
Medical specialities	18	22	19
Management	11	5	2
Primary care	10	10	5
Software	9	6	2
Medical informatics teaching	8	4	6
Education theory	8	9	1
Not related to informatics	7	6	2
CD-ROM	7	4	2
Evidence-based medicine	4	5	3
Telemedicine	4	2	2
Medical Literature	3	3	3
Decision support	3	2	1
Virtual Reality	2	-	-
Standards	1	2	1
Public health	1	-	1
Law	1	-	-
Emergency	1	1	1
Clinical guidelines	1	2	-
Information systems	1	5	-
Pharmacy	-	4	
Congresses	-	2	
Voice recognition	-	-	1

Table 7: Main topics observed in Medical Education Articles

Table 6 summarizes the last 10 years of evolution in the number of articles in searches using MeSH terms, compared with the non-specific CME term searches.

Term	1981 –1985	1986 –1990	1991 –1995	1996 –2000	Total
Internet	0	0	239	6104	6454
Intranet	10	587	2690	9090	12391
Software	3715	9112	13255	16728	44746
CD-ROM	1	52	295	471	819
e-mail	10	588	2712	9189	12505
Red TCP/IP	0	1	11	20	32
FTP protocol	0	1	20	42	62
On-line & computers	51	53	30	25	285
WWW	0	0	37	5098	5135
Virtual Reality	0	0	93	424	517
Bulletin Board System	0	6	29	27	62

Table 8: Number of Articles in Medline searches on specific terms (1981–October 2000)

Term	1996	1997	1998	1999	2000
Internet	442	689	1284	2222	1467
Software	3076	3683	3775	3824	2370
CD-ROM	106	102	106	93	64
e-mail	1419	1829	2136	2437	1368
Red TCP/IP	5	3	6	4	2
FTP Protocol	9	8	11	10	4
On-line & computers	11	5	5	3	1
WWW	108	218	1102	2230	1440
Virtual Reality	56	102	105	80	81

Table 9: Number of Articles in Medline searches on specific terms (1996-October 2000)

We also thoughtfully analysed those articles related to medical informatics and continuing medical education, excluding nursing, and classified them according to their main topics, as included in Medline abstracts (table 7).

In the second phase, we analysed the frequency of certain terms related to new information and communication technologies during the last 10 years. This is shown in table 8 and figure 2. Table 9 and figure 3 show detailed data about last 5 years.

Notice the use of the term Intranet in scientific articles before Internet technologies were known. Thus, 6 articles mentioned the term intranet from 1972 to 1976, even though it was used with a different meaning, and was applied to any local network. The term Internet appeared for the first time in 1991.

5 New perspectives

In the near future, it will not be easy to distinguish between continuing education resources and health information systems used by doctors in their daily practice of handling clinical history or generating reports. There is an increasing need to integrate screening or checking routines into information systems, which should advise physicians about the appropriate steps to follow. Even more, they could propose specific protocols used by their organization [Shea 96].

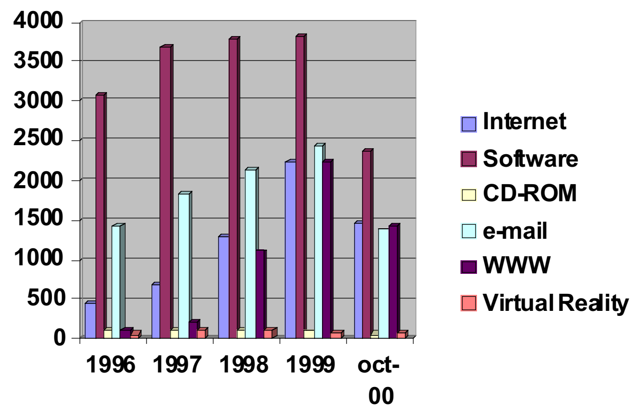


Fig. 3: Frequency of technical terms in scientific articles between 1996 and 2000

Therefore, the presence of integrated medical workstations will be common, an idea developed in the Columbia Presbyterian Medical Center, New York. They had a clinical information system (Web-CIS) with Internet access, access to the servers of the institution working as searching engines on the Internet and an Intranet. Consequently, it is easy to imagine future information systems capable of answering prompts such as “find cases that are similar to this patient's case and show me data about their diagnoses and treatments”. The answers will integrate information coming from the Intranet, the public Internet, or from certifying entities.

References

[Anonymous 00]
Anonymous. Barómetro on-line. *PC Actual*. 2000; 122: 289.
[Bravo Toledo 00]
Bravo Toledo R. *Medicina basada en la evidencia*. Lecture in the VIII Congreso Nacional de Informática Médica, Informed 2000. Toledo, Spain 4–6 October, 2000.
[Campbell/Johnson 99]
Campbell j.K., Johnson C. Trend spotting: fashions in medical education. *Br Med J* 1993; 318: 1272–1275.
[Ceusters/Prorec-Be 98]
Ceusters W. prorec-Be V.Z.W. Consensus management and promotion for high quality electronic patient records in Belgium. Paper presented in Prorec Seminars, Spain Seville. 1998.
[Coiera 98]
Coiera E. Medical informatics meets medical education. *Med J Aust* 1998; 168: 319–320.
[Collen 94]
Collen M.F. *The origins of informatics*. JAMIA 1994; 1: 91–107.
[du Boulay 00]
du boulay C. From CME to CPD: getting better at getting better? (Editorial). *Br Med J* 2000; 320: 393–394.
[Keravnou et al. 97]
Keravnou E., Garbay C., Baud R., Wyatt J. (eds.). Lecture Notes in Artificial Intelligence 1211: Artificial Intelligence in Medicine 6th Conference on artificial Intelligence in Medicine Europe, AIME’97, Grenoble, France, March 23–26, 1997: Proceedings. Springer, 1997
[Kidd et al. 93]
Kidd M.R., Connoley G.L., Cesnik B., McPhee W. What do medical students know about computers? *Med J Aust* 1993; 158: 283–284.
[LaPorte et al. 98]
LaPorte R.E., Marler E. Akazawa S. et al. The death of biomedical journals. *Br Med J* 1995; 310: 1387–1389.
[Ludvigsson 92]
Ludvigsson J. A curriculum should meet future demands. *Med Teach Editorial* 1999; 21: 127–129.
[Marimón 99]
Marimón S. (ed.). *La sanidad en la sociedad de la información*. Ediciones Díaz de Santos. Madrid. 1999.
[Moehr 00]
Moehr J. *Health Informatics theme – Overview*. http://healnet.mcmaster.ca/research.htm. Last access 23/10/00.
[Neame et al. 99]
neame R., Murphy B., Stitt F., Rake M. Universities without walls: Evolving paradigms in medical education. *Br Med J* 1999; 319: 1296.
[Nylenna et al. 99]
Nylenna M., Hjortdahl P., Aasland O.G. The use of Internet among Norvegian physicians. *Tidsskr Nor Laegeforen* 1999; 119: 4342–4344.
[Patel et al. 00]
Patel V., Kushniruk A., Arocha J., Kaufmann D., Moehr J., Lau F., Penn A., Grant A., Edmonds M. *Cognitive Evaluation of Emerging Health Information Technologies: A Proposed Integrative approach to HEALNet*. http://hiru.mcmaster.ca/nce/ Last access: 23/10/00.
[PubMed 00]
National Library of Medicine. PubMed. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed. Last access: 26/10/2000.
[Sandvik 99]
Sandvik H. Health information and interaction on the Internet: a survey of female urinary incontinence. *Br Med J* 1999; 319: 29–32.
[Shea et al. 96]
Shea S., Du Mouchel W., Bahamonde L. A meta-analysis of 16 randomised controlled trials to evaluate computer-based clinical reminders for preventative care in the ambulatory setting. *J Am Med Informatics Assoc* 1996; 3: 399–409.
[Snadden 99]
Snadden D. Portofolios – attempting to measure the immeasurable? *Med Educ* 1999; 33: 478–479.
[Verhoeven et al. 00]
Verhoeven A.A., Boerma E.J., Meyboom-de Jong B. Which literature retrieval is most effective for GPs? *Fam Pract* 2000; 17: 30–35.

to Contents, page 1

Automatic Three-dimensional Reconstruction of Radiological Images for Telediagnosis

Valentín Masero Vargas, Juan Miguel León Rojas, José Moreno del Pozo, Antmael Silva Luengo

In the field of Telemedicine, one of the most successful techniques has been remote diagnosis or Telediagnosis. To improve Telediagnosis, we have constructed an Information System for automatic three-dimensional reconstruction of radiological images. This system has improved the way in which we visualise and examine these radiological images. In this article, we describe the techniques that have been utilised for constructing 3D reconstructions. We also explain several segmentation techniques due to their importance during the 3D reconstruction process.

Keywords: Image Processing, Snakes, 3D Reconstruction, Computer Graphics, Medical Imaging, Telediagnosis, Telemedicine.

1 Introduction

1.1 Project goals

This work started with the project "Development of a 3D Database of Animal Organs (D3DBAO)", supported by the Spanish Intergovernmental Commission for Science and Technology (Comisión Interministerial de Ciencia y Tecnología, CICYT). This project is managed by the Minimally Invasive Surgery Centre (MISC) in collaboration with the Computer Science Department of the University of Extremadura (CSDUEX).

One of the main goals of our project was the automatic construction of both human and animal organs from radiological images. An additional objective was to use our software as a 3D visualisation environment and, later, as biomechanical simulation environment. All together, these will constitute a training system for surgeons based on virtual reality. These additional goals are related to another project called "Telesurgex (Minimally Invasive Surgery by Telecommunication)" supported by the European Community and the Junta de Extremadura (Consejería de Educación, Ciencia y Tecnología). Several researchers of the CSDUEX are collaborating in this project as well.

This article explains a system to automatically construct 3D models from radiological images. This system is still being improved.

While we were constructing this system, it was necessary to improve several techniques for automatically extracting 3D models from CT (Computed Tomography) and MR (Magnetic Resonance) images. Depending on the part of the anatomy needed for 3D reconstruction, it is better use CT or MR images. Thus, we worked with different radiological imaging techniques. For example, to reconstruct soft tissues it is better to use MR images, and to reconstruct a bone structure, it is better to utilise CT images.

2 Background

2.1 Three-dimensional reconstruction and visualisation of medical images

Thirty years ago, few radiologists were interested in three-dimensional imaging. They believed that this technique was not necessary because they could mentally assimilate two-dimensional data and interpret it in a three-dimensional manner.

Since then, three-dimensional imaging has become so advanced that image fidelity is sufficient for worldwide clinical use of three-dimensional imaging in dedicated academic medical centres.

By 1985, clinical application of medical imaging primarily focused on cranio-maxillo-facial surgery and was mainly limited to bone imaging on the basis of CT scans. At that time,

Valentín Masero Vargas is assistant professor at the Computer Science Department of the University of Extremadura. He first worked in the Minimally Invasive Surgery Centre. His main research interests are image processing, computer graphics and medical engineering mainly focused on telemedicine systems. vmasero@unex.es

Juan Miguel León Rojas is associate professor at the Applied Mathematics Department of the University of Extremadura. The subjects of his research are deviant logics, intelligent data analysis, decision support systems, machine learning and pattern recognition. <jmleon@unex.es>

José Moreno del Pozo is associate professor at the Computer Science Department of the University of Extremadura. He holds a Ph.D. in Physics. His line of research centres in image processing, computer graphics and pattern recognition. <josemore@unex.es>

Antmael Silva Luengo is assistant professor at the Computer Science Department of the University of Extremadura. His current research topics are gestural sign systems, machine learning and pattern recognition. <agua@unex.es>

to Contents, page 1

many radiologists feared being used as a kind of medical photographer merely to provide image data sets to surgeons [Hemmy et al. 94]. However, the complex segmentation process, especially in the case of soft tissue segmentation, clearly required radiologic skills, both to recognise and isolate the relevant structures in the two-dimensional cross-sectional scans. Later, three-dimensional imaging was introduced in other types of surgery with the help of other imaging modalities such as surface and volume rendering.

During the last decade, three-dimensional imaging has attracted more and more devotees mainly because of its realism and ability to interact with the 3D images. It allows visualisation from different views and several graphical transformations over the visualised three-dimensional objects. 3D visualisation has also been used for automatic detection of several diseases, such as cancer tumours, showing them in a different colour [Wyatt et al. 98]. In radiology, recent systems for distance diagnosis or telediagnosis have been developed, e.g. Hideyuki's system [Hideyuki et al. 98].

But, to be able to apply these three-dimensional imaging techniques, we must apply an efficient segmentation process beforehand.

2.2. The most widely used techniques for segmentation of medical images

Segmentation is the process of recognising objects within an image. As stated above, it is an indispensable step in most of the processes that are applied to medical imaging. For example, for measuring the volume of anatomical structures, for making three-dimensional reconstructions of medical images, or for making a good 3D rendering, it is important to develop an optimal segmentation.

It is important to emphasise that no general segmentation method exists which is able to isolate any object in any kind of image. That is, depending on the problem to solve and the given image, several segmentation methods are suitable. And, in most cases, no one method gives perfect results, so we need to adapt these techniques to our needs by using hybrid techniques.

The most widely used medical imaging segmentation techniques can be categorised into two classes:

1 feature thresholding methods,
2 region-based methods.

Among the region oriented methods, there are two subcategories: the region growing method and the texture analysis method. We do not mention the technique of cluster analysis which is basically a feature extracting technique and a multidimensional extension of the concept of thresholding.

Segmentation by thresholding has the advantage of being computationally simple and fast. It works well on objects with uniform intensity values which differ significantly from all other objects and the background. The thresholding fails, however, if the objects do not differ in intensity values but in some other property (e.g. texture).

Region-based methods assume the objects in the image to be a closed region. With this model of an object, two general approaches are possible: finding the border of the region (i.e. the edge oriented methods), or finding the interior of the region (i.e. the region based methods).

We distinguish two techniques to find the interior of a region. One technique is called texture analysis, whereby each pixel with its neighbourhood is classified individually as belonging to the region or not. The other technique is called region growing, an iterative method, whereby the image is processed several times. Each step of the iteration improves the result of the previous step. Problems discussed in this field are finding the seed points, the way of growth, the break of growth, and the split and merge of regions [Gonzalez/Woods 92].

Segmentation by edge detection methods is based on identifying significant intensity changes as edges of the objects in the image. The edge detection can be divided into the following steps. First, the image is preprocessed or conditioned, concentrating on brightness and contrast. Second, the first or second derivation of the signal is calculated. Finally, depending on the used derivation, the local maxima or the zero crossings are detected and marked as edge pixels [Gonzalez/Woods 92].

3 The Information System

The information system [Masero et al. 00] for automatic 3D reconstruction of radiological images is mainly formed by: a radiological device (RD), a workstation connected to the RD, a data server, an ATM telecommunications network, a Personal Computer (PC) connected to the system via Internet, and a computer program [Masero et al. 00]. In this article, the data server and the telecommunications system are called "Picture Archiving and Transmission System". The computer program is being finalised by the CSDUEX in collaboration with the MISC team.

The workstation allows us to provisionally store the radiological images and, optionally, process them for making the three-dimensional models. The user can make new 3D reconstructions remotely and visualise them on the screen of a remote PC connected to the system via Internet. From this PC several graphical operations can be applied to the 3D model, such as translations, rotations, scaling and, generally, all kinds of geometric transformations based on these operations. Therefore, the user can obtain new 3D images and study them better by using the options offered.

4 Methodology for 3D Reconstruction

Figure 1 shows the process for creating three-dimensional reconstructions. In this process, we have used a methodology based on image processing and computer graphics techniques.

Several computer graphics techniques, such as surface rendering and volume rendering, have been used.

Initially, all the techniques that appeared in the section, "The most widely used techniques for segmentation of medical images", were tested on several datasets, but no acceptable results were obtained. So, we realised that it is not possible to achieve good segmentation results utilising the same segmentation techniques for every kind of anatomical structure. Therefore, we applied different segmentation techniques, depending on the part of the body for which we wanted to make a 3D reconstruction. Finally, among all the applied techniques,

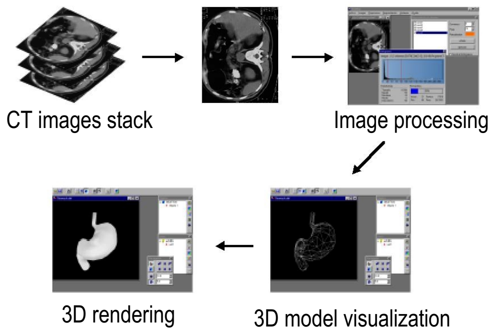


Fig. 1: 3D reconstruction process.

Snakes [Kass et al. 87][Honea et al. 97][Blake/Isard 98] was the method with the best results.

4.1. Snakes

Snakes were first described by Kass, Witkin and Terzopolous [Kass et al. 87]. A snake is a sequentially connected set of points, each located at one pixel in the image [Honea et al. 97]. The method of snakes is an energy minimization technique that uses information both on the object's shape and its image properties to determine the segmentation of the object. The snake has an energy, called external energy, that arises from the image characteristics at each snake's point location. The snake also has an internal energy, based on its shape.

This method requires that the snake be initialised to a path near the true object boundary. At each snake point, a set of nearby pixels is chosen as a neighbourhood in which the current snake point may move. During optimization, the lowest energy path is sought, and each snake point is moved, if necessary, to the path that has the minimum energy path. The process is repeated until no further movement of the snake is needed.

As result of applying an image operator to each snake point, we obtain the external energy at that pixel location. In this way we assign lowest energy to points that show good image qualities, and higher energy to other points. The good qualities are defined as those that usually happen at the border of the object we are examining (e.g. high intensity pixels). We can define the external energy as:

$$E_{ext} = -|\nabla I(x, y)|^2$$

The internal energy gives a smoothing constraint on the shape of the snake. This energy is formed by two terms. The first one causes the snake to behave as an elastic band that resists stretching. This first part can be written as:

$$E_{int1} = \sum_{i=1}^{n} |v_{i+1} - v_i|^2$$

n being the length of the vector and *v* a point of the vector where $v_i = (x_i, y_i)$.

The second part makes the snake behave as a flexible rod. It can be defined as:

$$E_{int2} = \sum_{i=1}^{n} |v_{i+1} - 2v_i + v_{i-1}|^2$$

All these terms are weighted to contribute in the appropriate influence to the whole snake energy.

$$E_{snake} = a \cdot E_{ext} + w_1 \cdot E_{int1} + w_2 \cdot E_{int2}$$

4.2. The whole 3D reconstruction process

In the system that has been built, the following process is needed to transform radiological images into three-dimensional models.

0. Setting environmental parameters, i.e. parameters associated with the radiological imaging device.
1. Acquisition of the radiological images. Always under conditions set on step 0.
2. Images preprocessing. Processes of brightness, contrast and histogram in order to eliminate noise and undesired information, that is, in order to have acceptable initial conditions.
3. Bidimensional segmentation: depending on the anatomical structure to delimit, we use a different segmentation technique or a hybrid one. For example, region growing or Snakes.
4. To select the regions obtained in the previous step and to connect them to form a three-dimensional model.
5. Three-dimensional models are made without additional operator interaction by propagating the 2D results to adjacent slices.
6. To export to appropriated graphical formats (ASC, 3DS and VRML) in order to later visualise and manage the 3D model in the developed visualisation software [Masero et al. 00]. The computer graphics techniques used have been Surface and Volume Rendering Techniques.

The software for visualising the 3D models has been developed using the visual programming environment Delphi 5® and the graphic library Open-GL®. It has been made for Windows 98® and Windows NT®. However, another version is being developed for Linux using C++.

5 Conclusions

We have designed an information system for obtaining three-dimensional reconstructions, remotely, from radiological images. This information system is based on a software which utilises several classic image processing methods [Gonzalez/Woods 92] and others more recent such as snakes [Kass et al. 87]. It also utilises a standard telecommunications network. To carry this out, a computer program has been developed. With this software, the physician can better study the zone of interest by studying it from several views. However, the main contribution of this work has been the improvement of the segmentation of radiological images by using snakes; and, therefore, the improvement of the 3D visualisation of the anatomical structures shown in the radiological images.

Our system, in addition to allowing the visualisation of stored images like the system described by Hideyuki [Hideyuki et al. 98], also permits interaction with previously stored 3D models. It also allows the creation of another 3D reconstruction from the images stored in the workstation or picture archiving system.

This software gives greater independence to physicians working in areas with limited resources. It allows the visualisation of 3D images of the anatomical structure of interest from any rural area. It also allows physicians who do not have a very powerful computer to control the remote three-dimensional images without having a workstation, which would be too expensive.

In the future, we would like to improve the visualisation of 3D imaging in order to improve diagnosis. We could also improve image segmentation that will allow us to see cancer polyps and other diseases more difficult to diagnose nowadays.

Acknowledgements

This work would not have been possible without the support of the Spanish Intergovernmental Commission for Science and Technology (Comisión Interministerial de Ciencia y Tecnología, CICYT) for the project "Development of a 3D Database of animal organs (D3DBAO)". We also wish to thank the European Community and the Junta de Extremadura (Consejería de Educación, Ciencia y Tecnología) for the "Telesurgex (Minimally Invasive Surgery by Telecommunication)" project support. These are two projects of the Minimally Invasive Surgery Centre (MISC) on which the Computer Science Department of the University of Extremadura (CSDUEX) is collaborating.

References

[Blake/Isard 98]
A. Blake, M. Isard: *Active Contours: The application of techniques from graphics, vision, control theory and statistics to visual tracking of shapes in motion*. London. Springer, 1998.

[Gonzalez/Woods 92]
R. Gonzalez, R. Woods: *Digital Image Processing*. New York. Addison-Wesley, 1992.

[Hemmy et al. 94]
D. C. Hemmy, F. W. Zonneveld, S. Lobregt, Z. Fukuta: *A decade of clinical three-dimensional imaging: a review.* Part I. Historical Development. Investigative Radiology. 1994; 29: 489–496.

[Hideyuki et al. 98]
B. Hideyuki et al.: A Teleradiology System with Realtime and E-mail-based Operating Modes using Synchronized Commands. In: H. U. Lemke, M. W. Vannier, K. Inamura, A. G. Farman eds. *Computer Assisted Radiology and Surgery. Proceedings of the twelfth International Symposium and Exhibition; 1998*. June 24–27 Tokyo. Elsevier Science BV, 1998: 449–453.

[Honea et al. 97]
D. M. Honea, Y. Ge, W. E. Snyder, P. F. Hemler, D. J. Vining: Lymph-Hode Segmentation Using Active Contours, in *Proceedings of SPIE Medical Imaging 1997*: Image Processing, vol. 3034, 1997.

[Kass et al. 87]
M. Kass, A. Witkin, D. Terzopoulos: Snakes: Active contour models. *Proc. 1st Int. Conf. Comp. Vision (1987)*, pp. 259–269.

[Masero et al. 00]
V. Masero, J. Moreno, F. Andres, A. Silva, J. Chambel, J. Uson: Telereconstruction from computerized tomography images for the distance teaching of medicine. *Journal of Telemedicine and Telecare*. 2000; Vol 6. Sup 2: 90–94.

[Masero et al. 00]
V. Masero, F. M. Sanchez, J. Uson: An integrated system of telemedicine for minimally invasive surgery. *Journal of Telemedicine and Telecare*. 2000; Vol 6. Sup 2: 88–89.

[Wyatt et al. 00]
C. L. Wyatt, Y. Ge, D. J. Vining: Automatic segmentation of the colon for virtual colonoscopy. *Computerized Medical Imaging and Graphics* 24 (2000) 1–9.

Artificial Intelligence in Medicine. Past and New Applications

José María Barreiro, José Crespo, Víctor Maojo

This article offers a general overview of the history of research in the field of Artificial Intelligence in Medicine. The first experiences were based on the use of cognitive models of medical reasoning, and gave rise to expert systems. Among these models that tried to emulate the language of physicians MYCIN was the best example. A second generation of systems were designed to fit better into the clinical routine environment. Finally, some examples of current research developed by the group are given such as: (1) a mobile system for emergency health care; (2) a system for accessing MEDLINE in Spanish, based on a medical vocabulary server; (3) a virtual system for training in trauma surgery. In these systems we use hybrid systems that combine Artificial Intelligence tools with Internet-based information management.

Keywords: Medical Informatics, Artificial Intelligence, Expert Systems, Internet, Databases, Emergencies, Virtual Surgery.

Introduction

Artificial Intelligence is a branch of Computer Science that begun at a meeting held at Dartmouth College in the United States in 1956. The original goal was to conceive intelligent machines that could replace human beings or perform tasks that could be considered similar to those of human beings. While initially the goal of making intelligent machines was an abstract one, it soon became clear that there was a possibility of creating intelligent systems for practical real-world applications.

After completing research into the cognitive issues of medical reasoning and computer-assisted diagnostic systems, researchers from universities like Stanford, MIT, Rutgers and Pittsburgh started to work in the 70s on what are known as expert systems, medical examples which include MYCIN, PIP, CASNET or INTERNIST.

Some years earlier, during the 60s, several groups of physicians at hospitals in Boston, Salt Lake City, in the USA, and Geneva (Switzerland) had proposed the use of early commercial mainframes, like UNIVAC and its successors, in clinical applications. This led, for example, to the early computer-assisted diagnostic systems and the first computerised clinical records. The above clinical applications were implemented at hospital computer centres by programmers, systems analysts and clinical physicians interested in their use. The goal was to create applications that could perform the calculations which occupied so much of physicians' time.

The following list provide an outline of the main methods used at that time in computer-assisted decision-making:

1. Decision analysis
2. Mathematical models of physiological processes
3. Databases
4. Clinical algorithms
5. Pattern-matching methods

Some researchers had been working on the study of reasoning processes required for issues of medical practice, such as, for example, clinical diagnostics. By the end of the 60s, they had found that diagnostics is an iterative, inference-based process of determining the type and circumstances of a disease by examination. Clinical reasoning was no longer viewed as an art, the result of a clinical eye or the intuition of the best specialists, and came to be considered a formal question that could be modelled. Its main components were the following: produce a first hand identification of the cause of the symptoms and signs of the patient, gather important information, select the best diagnostic tests and recommend treatment. This process is governed primarily by the hypothetical-deductive method, as a

José María Barreiro Sorrivas is M.D. and has a PhD in Computer Science. Associate Professor in the Artificial Intelligence Department, School of Computer Science, Polytechnic University of Madrid. His main research interests are Artificial Intelligence, genetic algorithms, artificial neural networks, knowledge-based models, medical informatics. He has edited several books and published more than 40 articles and communications in the field of Medical Informatics.

Victor Maojo García is M.D. and has a PhD in Computer Science. Associate Professor in the Artificial Intelligence Department, School of Computer Science, Polytechnic University of Madrid. Postdoctoral fellow at Georgia Tech and Harvard University-MIT. Member of several scientific societies and program committees of national and international conferences. Author of more than 50 articles, book chapters and abstracts. Evaluator of research projects for the European Union and for the Spanish R&D National Plan. He has obtained several prizes and grants from research agencies, scientific societies and industry

José Crespo del Arco has a PhD in Electrical Engineering from the Georgia Institute of Technology (USA). M.B.A. Associate Professor in the School of Computer Science, Polytechnic University of Madrid. His main research interests are: medical images, 3D visualization, Medical Informatics, algorithms and Artificial Intelligence.

to Contents, page 1

series of hypotheses are stated and refined by means of methods of deduction.

The research of Simon and Chase at Carnegie Mellon and that of other researchers, like Tversky, Kahnemann, Pauker or Anderson, laid the foundations that were to unveil the fundamentals of the cognitive processes of general and medical reasoning and for the use of this knowledge to build systems that reasoned similarly to physicians. They noticed that physicians used a series of rules of thumb or heuristics to reduce the search space required to achieve their goal (for example, diagnosis). Physicians estimated the probabilities of diseases by comparing diseases rather than using data from formal probabilistic statistics or decision analysis. The main heuristics were:

1. Heuristics of representativeness: the probability of a finding is estimated by comparison with population statistics.
2. Heuristics of availability: the probability of an event is estimated according to how easily it can be remembered.
3. Heuristics of anchoring and adjustment: physicians make an initial estimate, which is then adjusted depending on the characteristics of the patient under examination until they arrive at a final assessment.

Other processes involved would be, for example, generalisations and abstractions of facts discovered during the reasoning process, recognition of patterns to quickly state diagnostic hypotheses or the application of joint actions prepared beforehand.

These early computer-assisted diagnostic systems did not work well because physicians found it difficult to deal with probabilistic concepts. They discovered that the systems had to speak the same language as physicians and be able to explain their reasoning rather than act like black boxes. The earlier success of expert systems, like DENDRAL, in other areas proved that it was possible to capture the knowledge and reasoning methods of the best medical specialists.

MYCIN is the most widely known early medical expert system. It was designed for the treatment of infectious diseases (septicaemia and meningitis). It was an academic, rather than a clinical success, as it was not routinely used outside Stanford, where Shortliffe created it. Indeed, its creator says that it was used in only one hospital in Japan, because a head of department forced his team to do so. MYCIN was an IF… THEN… rule-based expert system, containing about 450 rules.

MYCIN also introduced a new method for handling uncertainty, which was created informally within the system. It consisted in the use of certainty factors associated with knowledge, which were later proven to be equivalent to Bayes' models.

MYCIN had a fairly novel interface for the time, although physicians did not find it easy to use. An example of an interactive consultation with MYCIN is given below:

```
1. Patient's name: (first-last
   ** FRED BRAUN
2. Sex:       (Questions 1-3 request
   ** M       background patient data.)
3. Age:
   ** 55
4. Are there any cultures for Fred Braun
   which may be related to the present
   illness, and from which organisms have
   been grown successfully in the
   microbiology laboratory?
   ** Y ----CULTURE-1-----
5. From what site was the specimen for
   CULTURE-1 taken?
   **BLOD =BLOOD (System corrects typing
   error.)
6. Please give the date and time when
   CULTURE-1 was obtained. (mo/da/yr time)
   ** JUN 20,1977 The first organism isolated
   from the blood culture of 20-JUN-77
   (CULTURE-1) will be referred to as:
   ----ORGANISM-1-----
7. Enter the laboratory-reported identity of
   ORGANISM-1:
   ** UNKNOWN (Typically identity is not yet
   known.)
8. The stain (gram or Ziedl-Neelsen acid-
   fast) of ORGANISM-1:
   ** NEG (Preliminary lab results give some
   clues.)
9. Is ORGANISM-1 rod or coccus (etc.):
   ** ROD
10. What is the form of the individual
    organisms (e.g. Iancet- shaped for
    cocci, fusifomm for rods, etc.)?
    ** FUSIFORM
…
```

The knowledge base could be modified so that engineers could introduce new knowledge, an idea that, due to scientific discovery, was an absolute must in Medicine. A module was also created to justify system reasoning, which showed the list of rules used by MYCIN.

The program, called inference engine, that handled the knowledge was successfully marketed and used to build expert systems in other areas. Paradoxically, the MYCIN model and spin-off tools were used to develop expert systems in areas far from Medicine, such as defence, banking or industry.

MYCIN was created as a question of academic research rather than in response to a real clinical problem. Therefore, it had the following problems:

(1) A consultation took longer than 15 minutes, time that a physician normally could not spend on just one consultation.
(2) MYCIN was more confident about its diagnoses than specialists, but its errors led to mistrust on the part of users.
(3) The explanations given by the system about its findings were simple and were only justified by the list of rules that had been fired.
(4) The interface was neither graphic nor easy to use.
(5) Physicians had to use MYCIN outside the clinical environment with the help of computer scientists.
(6) The knowledge contained in the system was heuristic-based but contained almost no deep information, that is, about the causes and processes underlying the disease.

To be able to replace physicians, programs were generally designed as absolute oracles. Even though the expert systems contained specialist knowledge in a field, they failed to model how people reason, including the use of common sense.

to Contents, page 1

Medical Informatics emerged as a separate discipline in the 80s, and the appearance of personal computers meant that physicians had access to applications for clinical, academic, research and management use. This led to the gradual emergence of a new paradigm: physicians should be able to access all the medical information and knowledge they needed from their computers. Therefore, artificial intelligence techniques combined with software engineering techniques create more efficient systems integrated into the medical environment.

The next generation of expert systems tried to capture knowledge in a way similar to the one physicians use in their reasoning strategies. MYCIN or INTERNIST, based on heuristics or shallow knowledge, could not capture the complexity of the knowledge of medical specialists. These new systems incorporated deep knowledge on the causes of a particular disease, including new reasoning capabilities, qualitative, causal, temporal, semi-quantitative models, etc.

Systems like ABEL, for analysing acid-base disorders, or ROUNDSMAN, with a library of knowledge composed of many articles, search for those most closely related to the patient in question, or KARDIO, which used an automatic learning system to elicit rules from the system of large databases, marked a change of approach from the early expert systems.

These systems were not designed to answer questions raised by an expert in a given domain. Nor were they intended to model the diagnostic process. They were primarily systems that gave specific advice or criticised certain medical decisions. New temporal and probabilistic models were also introduced that provided greater capabilities. In addition, graphical user interfaces were created, which made systems more user friendly. Issues like logical knowledge verification and more extensive validations were considered in more detail to assure their clinical impact outside the laboratory in which they were created.

They were only partially successful as well because they had similar drawbacks like the 1st generation, as well as one in particular: physicians generally had neither the time nor the patience to enter into a prolonged dialogue with the computer, into which they also had to manually enter patient information.

90s

Twenty years after the first expert systems appeared, they were still not clinically successful in routine medical practice. Only systems like INTERNIST (now called Quick Medical Reference or QMR), for example, had been commercialised and used in a host of health system environments. As a result, new artificial intelligence-based models emerged using different concepts founded on new symbolic methods and, also, on connectionist models. Of these, artificial neural networks and case-based reasoning were prominent.

As of the 90s, new medical informatics systems based on the use of artificial neural networks appeared. They have been integrated into commercial systems and used, for example, in radiological diagnosis (for example, early cancer detection), in the classification of electrocardiograms and electroencephalograms, or in ICU systems.

Artificial neural networks can also be used as a method of data mining for eliciting knowledge from data warehouses or large databases (for example, complex clinical records).

Artificial neural networks were not as successful as originally thought, due to several factors. Problems include the number of cases required to train and validate the system, the epidemiological mistakes made in selecting these cases, the black box concept of these systems and their own errors. Thus they are trusted by clinical physicians (although they were often more correct than the specialists with whom they were compared).

The other model, case-based reasoning, considers the physician's diagnose of a patient in comparison with earlier cases. Thus, when faced with a new patient, the system searches for a similar patient in a case library. It makes a comparison, establishes the differences in order to reach a diagnosis and recommends a series of actions. This information is then stored in the library and the system operates accordingly.

The problem with this technique is that very similar cases are seldom found in Medicine. Moreover, the fundamental cognitive question is whether or not physicians actually use a combination of several reasoning strategies, which can be related to several of the techniques discussed above (heuristic search, pattern recognition, qualitative and causal reasoning, case-based reasoning). Therefore, one technique alone cannot capture the complexity of medical reasoning.

A combination of all these techniques would be an ideal formula for building an artificial intelligence system in medicine. Examples of these have already appeared, like PERFEX, a computer-aided SPECT image diagnostic system that combines a rule-based symbolic model with an artificial neural network connectionist model.

Present and future research

A host of clinically unsuccessful programs based on artificial intelligence techniques have been built over the last 25 years. Indeed, many were built merely as research experiments, and were not designed as solidly as, for example, the clinical trials run on new drugs. But some systems have been used clinically and commercially. Not only are these systems not public, their use has not been publicised (for example, some US medical insurance companies have expert systems for evaluating the quality and costs of health care).

In recent years, increasingly small systems have been designed for integration into hospital information systems (called HIS), like what are known as Medical Logical Modules, based on ARDEN Syntax. This syntax, similar to PASCAL, can be used to specify small computer-aided decision-making systems that can be easily integrated into more complex systems.

The rise of what is known as evidence-based medicine has promoted the development of programs that are based on negotiated knowledge, a combination of scientific literature and the opinion of expert panels, expressed as clinical practice guidelines and protocols. A host of software tools has been built with languages suitable for specifying protocols, which interact directly with clinical records to extract the data required to support decision-making. One of the models most commonly used for representation is the flow diagram, with state, action and

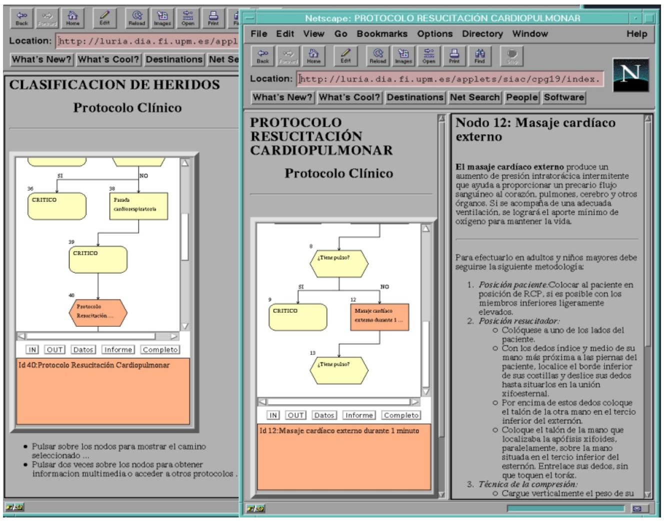


Figure 1: Emergency protocol management (version 1 in JAVA)

decision nodes, as illustrated in Figure 1, showing a system developed by our laboratory for managing emergency protocols.

Using the standard interfaces of common navigators and systems based on the use of CORBA or JAVA, new technologies can be used across the World Wide Web to develop small applications or components. These components are designed for very specific functions, which interchange messages and operate within a clinical workstation. Thus, computer system developers can easily implement small applications that can be reused in other domains or interchanged with other groups.

One essential requirement for the interchange of information is a common vocabulary which can be understood by all applications. This is particularly difficult in Medicine, where one concept can have different names in a specific same language. Therefore, an idea emerged, again in the USA, of a unified vocabulary which integrated all the existing medical vocabularies and could act as the link between different information systems from a semantic viewpoint. This vocabulary, called Unified Medical Language System, UMLS, is particularly useful for vocabulary servers.

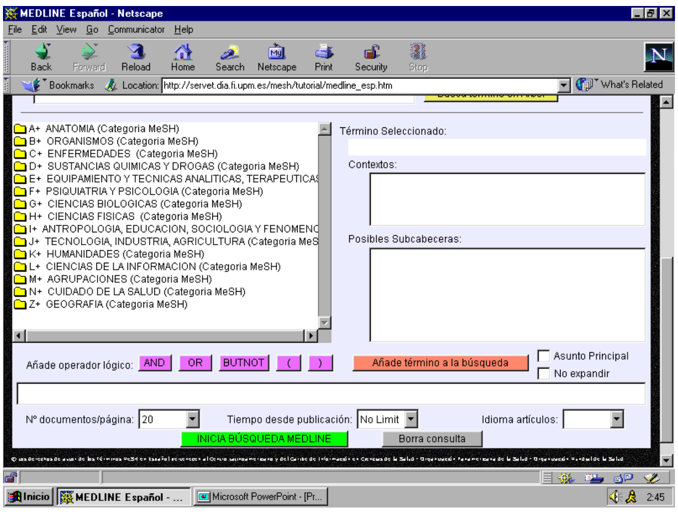


Figure 2: Access to MEDLINE in Spanish.

Our group has developed a vocabulary server, based on UMLS and other well defined medical vocabularies, designed to exchange information between heterogeneous systems. This server can be used over the Internet to identify terms or codes from different remote medical systems by means of Artificial Intelligence techniques. It has also been used to develop a system implemented in JAVA to access MEDLINE using Spanish terms, which is shown in Figure 2.

The nucleus of medical practice is the patient. Similarly, the nucleus of medical informatics is the computerised clinical record. This document contains the data required by physicians to assess, diagnose and treat a patient. Physicians should have systems related to clinical records as aids for managing the data, information and knowledge required for health care. Systems are needed now and in the near future to help to search and select the necessary information within the World Wide Web.

In the USA, the NII (National Information Infrastructure) initiative and the forthcoming generations of Internet, with Gigabits/sec transmission speeds, will soon provide real-time access to information contained in graphics, video, sound and the routine creation of telemedicine applications. This will make it possible to transmit all sorts of medical images, virtual reality applications, videoconferences, etc., which will radically change the face of medical practice.

Our group is working on a joint project, with the Ministry of Defence's General Health Inspection Unit, concerning the transmission of images for use in distance surgery planning and in virtual head injury surgery systems. The goal of this project is to help surgeons perform real operations and provide a virtual education system for training surgeons. Even though, fortunately, Spanish military physicians are unlikely to be involved in many wars, these systems can help simulate real situations and hence improve the training of these surgeons. Since it takes 10 to 15 years to train an expert surgeon, the use of these intelligent teaching systems can reduce this span of time and improve the response for situations that are unusual in peacetime. A screen of the prototype is shown in Figure 3.

With this and other similar projects, patients already have access to this new information, which heralds a genuine revolution in medicine. No longer merely a subject for study, patients will become agents that access, exchange and generate information. Therefore, systems have to be built not only to assist health workers but also patients. These include intelligent systems, like agents that navigate, search and filter important information over the Internet. They are called guardian angels, which are systems integrated into each patient's clinical records that know their background and preferences and can rule out any decision that could otherwise compromise their quality of life.

In the United States, new genetic clinical records are being created, in which diverse genetic information can be efficiently added. Therefore, the results of technologies like biochips are likely to have an immediate impact on clinical practice even in

to Contents, page 1

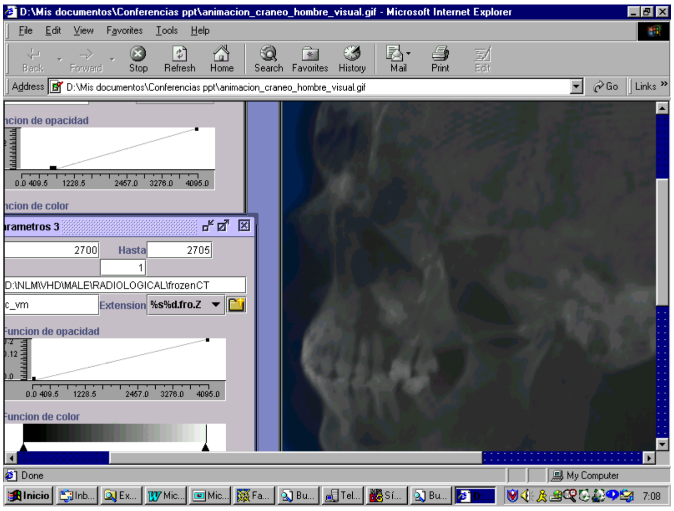


Figure 3: 3D navigation system for head injuries

primary care. Thus, it will also be possible to analyse risk factors or regional prevention campaigns. They will probably be commercialised by 2002.

The latest generation of mobile telephony technology, like WAP and, especially, UMTS and their immediate successors, as well as its integration into agenda-sized computers (personal digital assistants or PDAs) with improved screens and architectures, will give health care workers and patients multiple, mobile, universal and rapid access to remote computer-aided decision-making systems. This includes access, for example, to systems integrated with clinical records or radiological images from any part of the world with no access to cable networks.

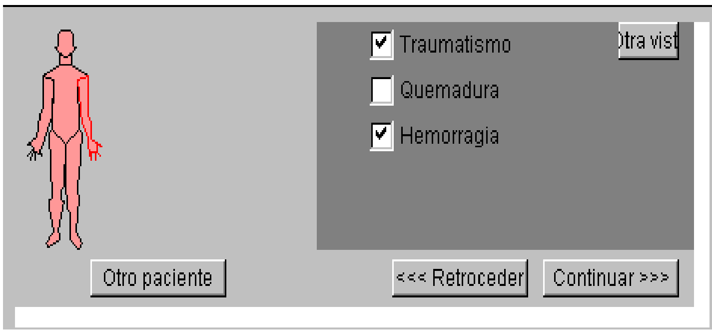


Figure 4: Emergency aid prototype system for PDAs

The European Commission is supporting this future technology by carrying out research projects that could be concluded within the 2002/04 period and then marketed.

Figure 4 is an example f a screen of a system implemented by our group jointly with physicians from the EMAT, at the Gómez-Ulla military hospital of Madrid and the Ministry of Defence's General Health Inspection Unit. It is designed to handle victims in two sorts of emergency situations: military emergencies (victims of combat in wars) or civil emergencies (common medical emergencies or disaster situations). This system is designed to be for use in portable computers and future PDAs, which are connected to computers based at reference hospitals, and can be adapted for use within UMTS networks.

References

Bleich, H.L.: (1969). "Computer Evaluation of Acid-Base Disorders". *Journal of Clinical Investigation* ,48,1689–1696.

Buchanan, B. and Shortliffe, E.H.: (1984) "*Rule-Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project*". Addison-Wesley, New York, USA.

Clancey, W. and Shortliffe, E.H. (Eds) (1984): *Readings in Medical Artificial Intelligence: The First Decade*. Addison Wesley. Reading, MA.

Evans, D., and Patel, V. (Eds) (1989). *Cognitive Science in Medicine. Biomedical Modeling*. The MIT Press. Boston, MA.

Gorry, G.A., Kassirer, J.P., Essig, A., and Schwartz, WB (1973): Decision Analysis as the Basis for Computer-Aided Management of Acute Renal Failure, *American Journal of Medicine*, 55, 473–84.

Greenes, R.A. and Shortliffe, E.H.(1990): "Medical Informatics: An Emerging Academic Discipline and Institutional Priority". *JAMA*, 263, 1114–1120.

Kassirer, J. and Kopelman, R. (1991): *Learning Clinical Reasoning*. Williams and Wilkins. Baltimore, USA.

Miller, P. (Edit) (1988). *Selected Topics in Medical Artificial Intelligence*. Springer Verlag. New York, USA.

Reggia, J. and Tuhrim, S. (1985): *Computer-Assisted Medical Decision Making*. Springer Verlag. New York.

Schwartz, W.B. (1970). "Medicine and the Computer: The Promise and Problems of Change". *New England Journal of Medicine*, 283, 1257–1264.

Shortiffe, E. and Perreault, L. (1990): *Medical Informatics. Computer Applications in Health Care*. Addison Wesley. Reading, MA.

Sox, H., Blatt, M., Higgins, M. and Marton, K. (1988): *Medical Decision Making*. Butterworths. Boston, USA.

Szolovits. P.(Ed.) (1982): *Artificial Intelligence in Medicine*. Westview Press. Boulder, CO.

Van Bemmel, J. and Musen, M. (1997): *Handbook of Medical Informatics*. Springer-Verlag. New York, USA.

Weiss, S.M., Kulikowsky, C.A., Amarel, S., Safir, A. (1978). A Model-Based Method for Computer-Aided Medical Decision Making, *Artificial Intelligence*, 11, 145–72.

to Contents, page 1

A Case Based Reasoning Decision Support System for use in Medicine

Joël Colloc and Laïd Bouzidi

Our DSS integrates different kinds of decision models and uses them to deal with the successive clinical decision steps: diagnosis, prognosis and therapy in the complex medical field. The decision process includes patient data, experts' knowledge, statistical and epidemiological data and previously stored experience. The Case Based Reasoning (CBR) approach is used to store and retrieve the clinical cases. This approach is viable in other complex fields that require a high level of expertise such as engineering design, computer design.

Keywords: Decision Support System, Case Based Reasoning, Multi-Agent system, Machine Learning, Clinical Decision, Knowledge Base, Epidemiology

1 Introduction: the clinical decision context

A decision problem can be defined as a choice between several options in order to achieve a goal as efficiently as possible. Most of the time, the physician's ability is beyond the computer's capability and thus, trying to supply a decision support system to aid him or her is worthless. The best example is diagnosis. However, medical activity could be defined as a *decision chain* (more exactly a decision network) that involves the following steps: the diagnosis, the prognosis, the therapeutic and the treatment follow-up.

Some of these steps become more and more complex and an aid tool will be useful, especially during the prognosis and therapy stages.

Artificial intelligence scientists were very interested in the cognitive nature of the physician diagnosis activity. So, many applications have been developed in this domain. We have to admit that most of them are unused, because they do not bring any actual enhancement. Paradoxically, many physicians are interested in information systems and some of them have developed small or sometimes more sophisticated Decision Support Systems (DSS).

1.1 Medicine complexity

Medicine is a science based on the human organism observation. Diseases are dynamic processes, but clinical syndromes are only pathologic snapshots, corresponding to evolution stages. Some diseases progress by steps in a linear way.

For example multiple sclerosis displays successive evolution steps. Others progress in a cyclic manner such as the duodenal ulcer or the herpes infection. Therefore, time is a major factor when describing pathologies [Summons et al. 98].

The effect of an inappropriate therapy or another underlying disease can interfere and give unusual clinical aspects.

1.2 Integrating quantitative and qualitative decision methods

Usually, decision support models in health care can be grouped into two main categories: quantitative and qualitative models.

Quantitative models are based on statistical methods and try to assess the disease occurrence probability for a given patient belonging to a population, when he or she shows some clinical signs or symptoms.

Qualitative models rely on expert knowledge and symbolic reasoning methods, which handle Boolean logic rules. When using qualitative methods one needs to cope with issues such as the experts' knowledge acquisition bottleneck, heuristic detection and the knowledge base implementation.

The models can also be grouped according to the learning method employed. Two types of learning are used: *supervised learning* in which, for each case of a training set, the correct solution is provided to the system by one or several experts, and *unsupervised learning* where the system must automatically determine which feature subset, or cluster of a set of features is relevant to characterize a given identified situation (i.e., a dis-

Joël Colloc is associate professor in computer sciences in the Jean Moulin (Lyon 3) University. His main research topics are DSS and knowledge bases in medicine, CBR and multi-agent systems, nervous system modelling using reactive MAS and genetic approaches. He holds a Ph.D. (1990) in computer sciences by the National Institute of Applied Sciences (INSA) of Lyon and a MD (1985) by the medical faculty of the Claude Bernard University (Lyon 1).

Laïd Bouzidi is associate professor in computer sciences in the Jean Moulin (Lyon 3) University. He first worked on documentary computing and then on decision support systems and Web information systems. He holds a Computer Sciences Ph.D. (1989) by the Claude Bernard University (Lyon 1) and an AI undergraduate degree from Aix Marseille 2 Luminy University.

Vous trouverez une version française de cet article sous <http://www.access.ch/sinfo/pages/issues/n011/no011.html>.

	Supervised learning	Unsupervised learning
Quantitative models	Neural networks Fuzzy sets Bayesian models	Neural networks Genetic algorithms Case-based reasoning
Qualitative models	Semantic models Object and frames Logic rules Decision trees	Case-based reasoning

Table 1: A decision model classification

ease) [Van Bemmel et al. 97]. Table 1 displays a Caroll diagram that represents a classification of the main decision models.

1.3 The Case Based Reasoning Model

The Case Based Reasoning (CBR) approach can be described as a third intermediate model, because it makes use of concepts coming from both qualitative and quantitative methods. On one hand, the quantitative methods provide the CBR approach with tools to automatically select the relevant features and to achieve the previously solved case indexing. On the other hand, the qualitative methods offer means to represent the experience gained in solving previous problems. Knowledge models have indeed the necessary semantic capabilities and thus are well suited to describe the case environment and circumstances. We briefly present a CBR model to store the clinical experience in the third section of this paper.

2 The medical decision process features

2.1 Components and stages of a standard decision process

According to Simon [Simon 77], the decision process is composed of four successive stages: 1. Intelligence: collecting static and dynamic data; 2. Design: designing scenarios; 3. Choice: choosing the more relevant scenario; 4. Implementation: evaluating the issues and if necessary providing a strategy update. Data collection is mainly supported by data bases and mathematical or statistical model bases [Pomerol et al. 93]. Dynamic information is provided by the decision-makers and is stored in knowledge bases. The decision bases are designed and implemented with the help of subsystem 1 and subsystem 2 (Figure 1). The chosen decision evaluation is based on the consecutive action impact analysis.

The DSS is composed of four subsystems:

- *Subsystem 1* (intelligence) is used to collect static data and is mainly based on databases, mathematical and statistical models and their corresponding management systems. Therefore, qualitative and quantitative data are extracted and then used by treatments according to the decision-maker. Relevant information is provided to the decision-maker in various formats: feature tables, scenarios, forecasting simulations, dashboards etc. Collectively this information is called *result Info* in Figure 1.
- *Subsystem 2* (design) collects dynamic information based on experienced decision-maker's knowledge and know-how. The domain knowledge is structured into chunks that are distributed according to a competence hierarchy or lattice. Subsystem execution is controlled by an inference engine (IE) but is supervised by the decision-maker. The result is an expert knowledge base (KB).
- *Subsystem 3* (choice) is concerned with designing and generating the decision base. It helps the decision-maker build a set of relevant decisions. This operation is based on information provided by a case base, which provides knowledge about similar previous experiences, and subsystems 1 and 2. The case base memorises the context and the previous reasoning processes used to solve (or fail to solve) previous problems. The output of Subsystem 3 is a decision matrix called the *decision base* in Figure 1.
- *Subsystem 4* (implementation) is used to choose the more appropriate decision between those selected during the

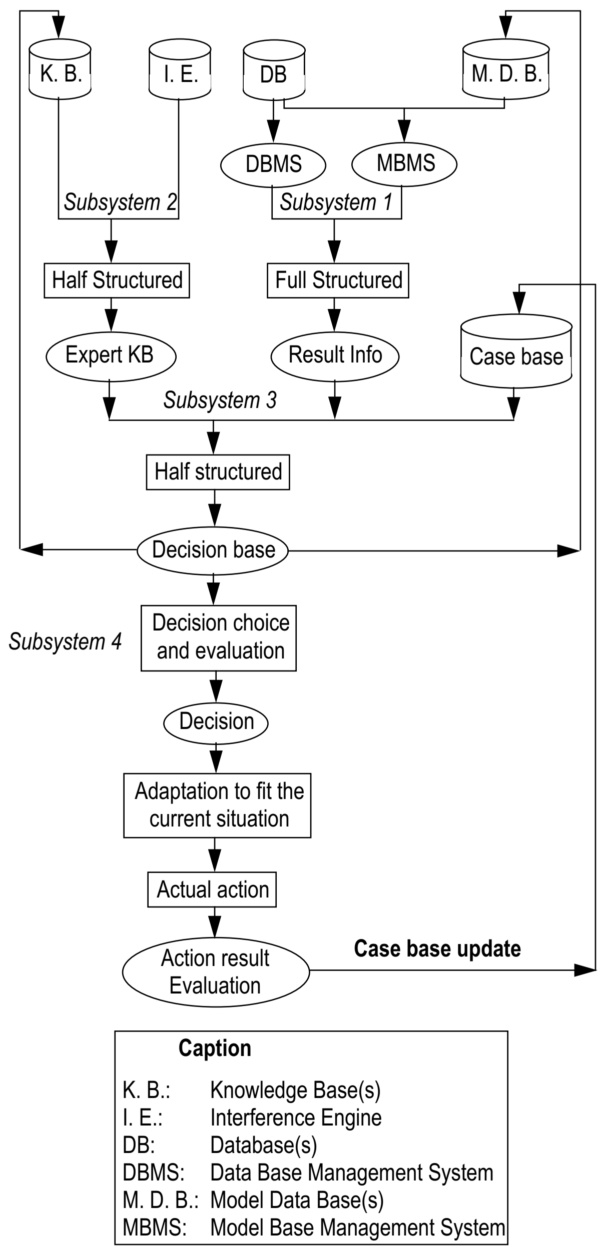


Fig. 1: The decision support system architecture

previous stage and to evaluate the impact of the applied decision. During a first step, the decision-maker browses and analyses the weight of each potential decision. Next, he has to list the relevant actions to achieve the selected decision (*strategic level*, *operational level* and *control level* actions).

Finally the decision-maker will be invited to evaluate the results and the impact of the selected decision in terms of facts, events, action consequences and to enrich the case base with the gained experience. This operation is depicted on Figure 1 by the *case base update* labelled arrow. Thus subsystem 4 closes the loop of the system's knowledge acquisition process.

2.2 A clinical decision model

The physician must predict the disease evolution, and know the previous therapies and their results in order to choose an appropriate treatment to improve the patient's health. One important aspect of a physician's knowledge is their related clinical experiences with similar cases (the essence of the case based reasoning approach).

We provide a definition and a representation for diagnosis, prognosis, therapy and therapeutic follow up.

The diagnosis is based on medical and surgical semiology: the clinical signs and events, which guide the physician to identify the disease. Semiology is a science that studies the clinical sign's nature, and how their combinations and evolutions define clinical pictures. Experienced physicians often memorized these decision schemes, made them unconscious and, thus, difficult to formalize. During the *positive diagnosis step (D+)*, the physician considers current complaints, the medical history, and the genetic and social background of the patient. Through examination of the patient, the physician searches for the clinical signs that we classify into three different types: pathognomonic signs, evocative sign and accessory signs. During the positive diagnosis step, the physician is making hypotheses relevant to the patient's clinical state. During the *differential diagnosis step (D≠)*, the physician is searching for the existence or the lack of specific signs in order to eliminate those which are not relevant from the previously elaborated hypotheses. The *aetiological diagnosis (Det)* is the correct result of the diagnosis process. It explains the disease appearance, manifestations and evolution causes. Sometimes, the physician is not able to find the disease aetiology *(D?)*, because the clinical picture is very unusual and he must act very quickly to avoid the patient's death. So, he must cure the main syndromes to get time to do further investigations. The diagnosis information will be useful to the prognosis as well.

The prognosis stage is often ignored in medical DSS. It is nevertheless essential because it allows the system to fix the therapeutic goal. The prognosis is a difficult task in the course of which the physician tries to predict the patients clinical state evolution and the probable outcome (healing, stabilizing, and death). The necessary knowledge and data derives from many sources: the aetiological diagnosis, the patient general clinical state (the genetic and social background, the psychological factors, the physical characteristics, the past medical history), the stage of the disease; the qualitative knowledge concerning well known clinical scenarios representing likely evolutions, as well as the quantitative knowledge provided by epidemiological studies based on statistical data. Perhaps, the most relevant knowledge comes from observation of some similar cases, which leads the physician to think that the evolution of the present disease will be similar. A CBR approach is used to memorize disease clinical evolutions.

The therapy step: The physician must know the available therapies, the indications for administering drugs, pharmacology data, contraindications, drug interactions, drug toxicity and

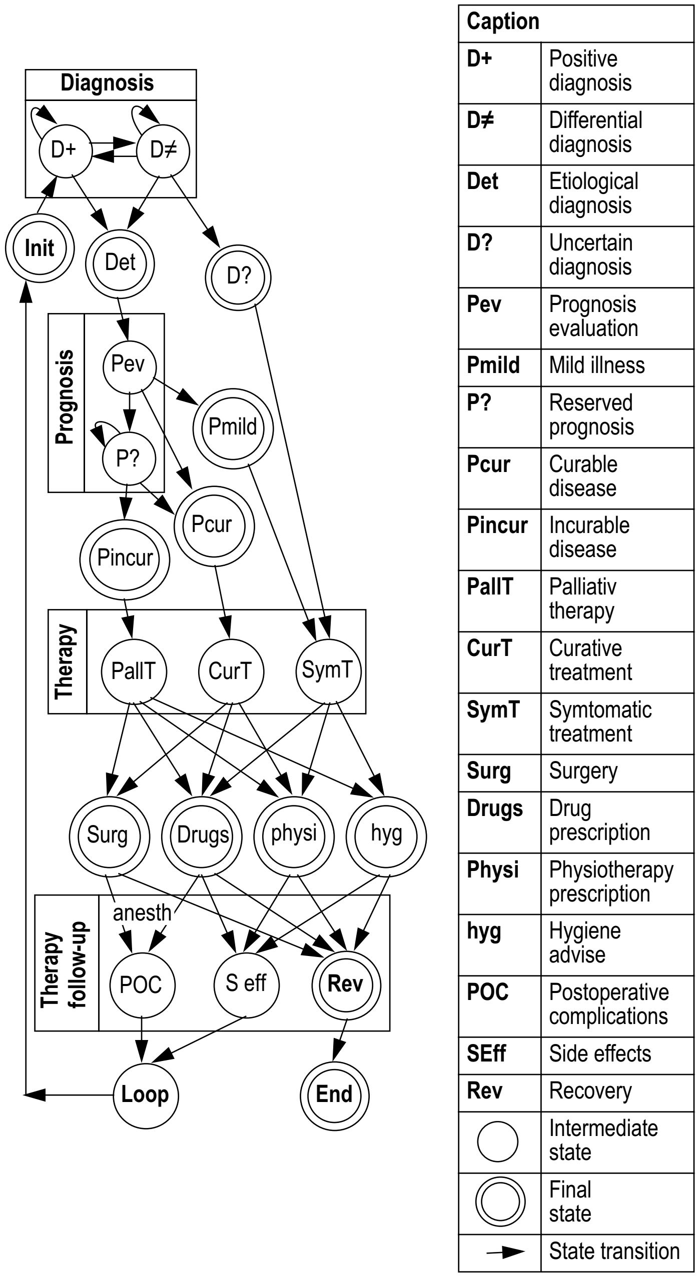


Fig. 2: The clinical decision system cycle

to Contents, page 1

adverse effects. The physician has previous prescription experience that allows him to predict the effect of a therapy strategy in similar cases. A therapy is prescribed to achieve a goal. A curative treatment (CurT) aims at totally curing the patient or consolidating his clinical state (to stop the disease evolution). For example: an antibiotic prescription to cure tonsillitis. A preventive treatment (PrvT) is prescribed to prevent a serious disease occurring, as in the case of vaccination or antibiotic prescription before a septic surgical operation. A Symptomatic treatment (SymT) cures disagreeable effects and functional manifestations of a disease, for example a headache. A Palliative treatment (PallT) is prescribed when the prognosis is hopeless and the aim is no longer to cure but to procure more comfort for the patient.

During the therapy follow-up step, the physician must watch over the patient's clinical state and the disease evolution towards recovery. He must diagnose the drug side effects, the postoperative complications or the appearance of an iatrogenic disease.

The whole decision process is represented by a finite state automaton, depicted in Figure 2, which triggers in turn each clinical module and thus coordinates them.

3 The Medical experience: storing and retrieving cases

Case-Based Reasoning (CBR) is a powerful concept that provides an analogic reasoning mode in problem solving [Aamodt/Plaza 94]. This capability allows CBR to express previous medical experience knowledge and thus to use it to enhance the diagnosis, prognosis, therapeutic and patient follow-up. The basis of CBR is that, by comparing new cases with previously stored, indexed clinical cases, and retrieving those cases that are similar, we can apply the corresponding decision and actions to the present patient, expecting that what was good one time, will be good several times [Gupta 94] [23]. The CBR approach includes the appropriate steps to deal with analogic reasoning.

Two main functions are provided: case storing through the "*new case indexation module*" and "*case retrieval*" handled by the so-called module in Figure 3. These two complementary features implement the CBR cycle.

The case base contains patient cases composed of diagnosis, prognosis and therapeutic facts. Figure 3 describes the main steps of the CBR cycle.

During the case retrieval stage, the CBR module computes structural similarities between the composite objects representing previously stored cases and the new clinical case under consideration.

A decomposition process of the case composite object provides sub-objects representing the following features: the problem definition and goal (PG), the environment representation (E), the reasoning protocol (RP), the applied decision (D), the necessary actions (A), and the actual result (Rs). During the case indexation stage, the new object case is instantiated and provided with diagnosis, prognosis and therapeutic components. The user must supply information that concerns the case

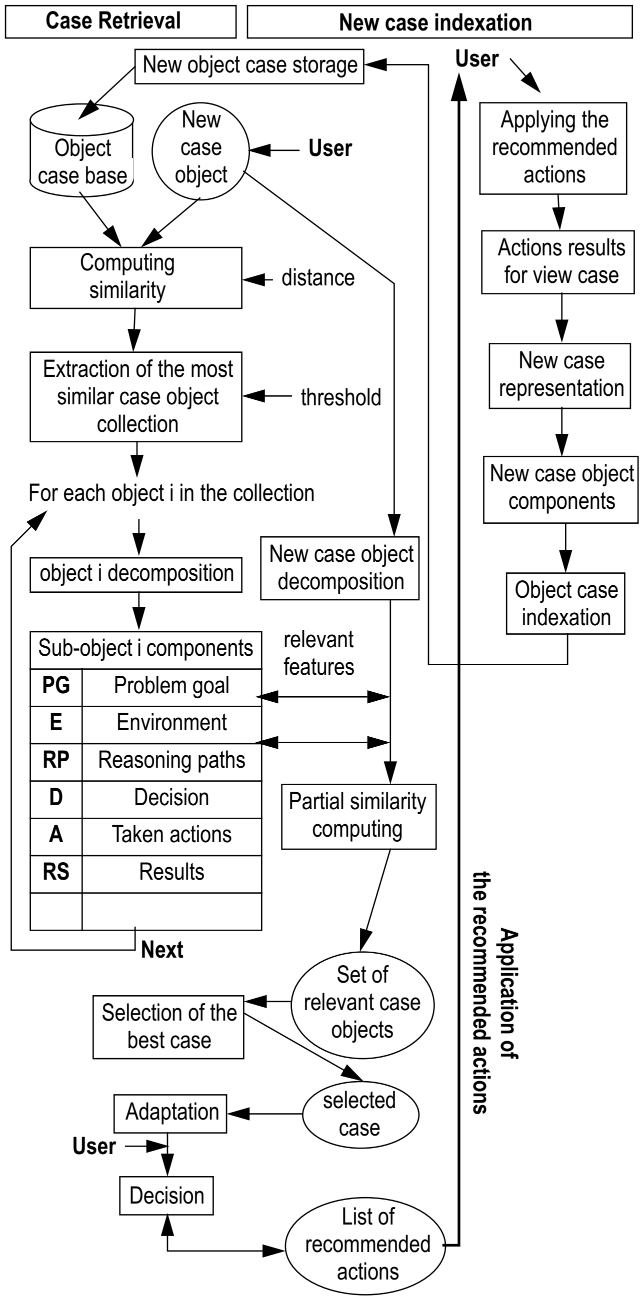


Fig. 3: The CBR cycle

features and circumstances {PG, E, RP, D, A, Rs}. Then the new case is indexed and stored in the case base. The process is depicted on the right side of Figure 3.

An application of the model to epilepsy diagnosis and therapy is available in [Colloc/Bouzidi 00].

Indexation of cases relies on a distance computation. Different distance models can be used to sort the cases: such as fuzzy logic [15] and the theory of evidence [Schuster et al. 97]. However, this topic is not within the scope of this paper and is therefore not further described.

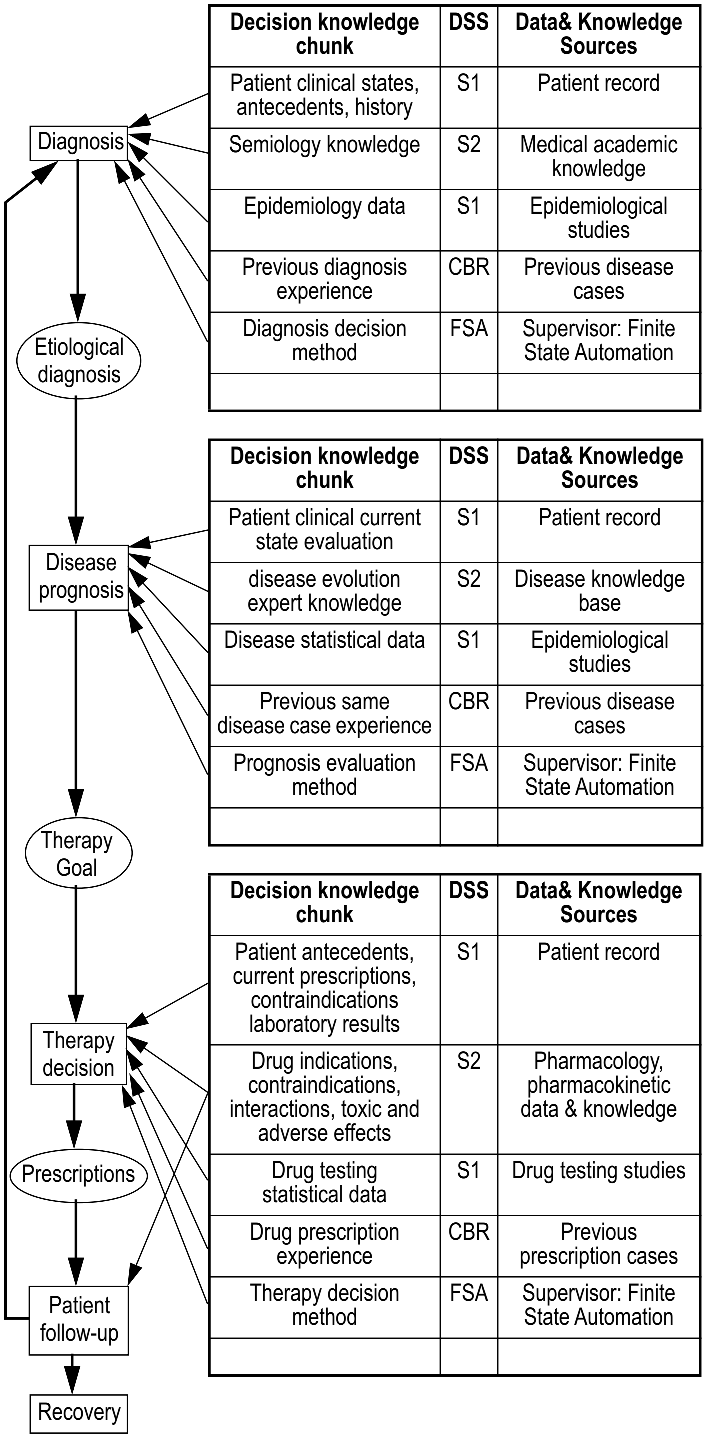


Fig. 4: The clinical DSS architecture and sources

4 A decision support system framework

4.1 The clinical decision process supervision

The purpose of our system is to integrate quantitative with qualitative decision methods. Most of the existing systems are based on only one of these two kinds of methods. To achieve this goal, some authors have proposed a flexible architecture which uses a Multi-Agent System (MAS) [Wiederhold 92] based on a communication language like KQML [Finin et al. 93] and a negotiation protocol like the contract net protocol [Jennings 93].

A MAS is an interesting solution to build such a decision support system but it is much more difficult to implement and to control. Our system uses a Finite State Automaton (FSA) which represents the states and transitions depicted in Figure 2 and triggers the appropriate procedures. This approach is similar to a supervised system [Schuster et al. 97]. The FSA describes the module interactions but it does not represent the data and knowledge flows. Some states are intermediate states and others represent a decision step final state.

4.2 Integrating the data and the knowledge sources in a decision cycle

For each clinical decision step, data and knowledge sources are necessary. Some of the data are provided through the interface by the user, others are stored in databases or knowledge bases. These information sources are displayed on Figure 1 in the subsystems 1, 2 and 3.

Figure 4 depicts the required knowledge chunks to achieve the decision, with the corresponding decision subsystem and the data or knowledge sources.

The small arrows in Figure 4 represent the relationship between the knowledge chunk and the decision step. The big arrows show the decision step succession and their results. Different data and knowledge sources are represented with different models but they must share a common interface. One solution is to use a multiagent system with an ontology that allows the DSS to share the same terminology. The agent is used to encapsulate the DSS module and to manage the communication and the negotiation between the agents in the system. Another solution is to use an object oriented approach to encapsulate the DSS modules and thus to inherit a general interface which allows them to communicate between each other.

The CBR module achieves transaction and clinical case recording. It must index and store the already-solved clinical cases coming with the knowledge chunks that were used to provide a solution. The environment consisting of the facts, events and especially the results, must also be stored. The main CBR drawback is that at the beginning of system use, the case base is void or contains very few cases. This gives no chance to find a case close to the one under consideration.

5 Conclusion

We showed that different kinds of DSS models, data and knowledge are complementary, and that they may all be useful to the determination of an appropriate decision in a complex domain like medicine. We have presented a framework to cope with the different decision paradigms integration. The system's supervision is managed by a finite state automaton that triggers queries to the appropriate database and knowledge base. We don't believe that a DSS can be implemented by a stand-alone CBR. This would require actual cases from the human being experience (patient records) be represented and stored from scratch; which is very difficult. We prefer an integrated approach involving knowledge and databases, which allows the initial priming and the ongoing enrichment of the case base. This process actually achieves a learning system. All these reasons justify the implementation of an integrated approach involving the different sources of information: KB, DB and CBR, to deal with the decision process. The coordination is done through a multi-agent system. A supervisor agent is implemented by a FSA that triggers the appropriate modules at each stage of the decision. It can be achieved with a syntax analyser like Yacc and the resulting C function is incorporated in a base class and then inherited by Java or C++ object classes. The main advantage of this approach is the system modularity. Therefore, the system can be built up incrementally. This increases system flexibility and feasibility.

References

[Aamodt/Plaza 94]
A. Aamodt, E. Plaza: Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches, *AI Communications*, IOS Press, 1994, Vol. 7: 1, pp. 39–59.

[Colloc/Bouzidi 00]
J. Colloc, L. Bouzidi: A Framework for Clinical Decision Making and Medical Experience Storing, in *Proceedings of ECIS'2000*, Vienna, Austria, 2000, pp. 1245–1252.

[Finin et al. 93]
T. Finin, J. Weber, G. Wiederhold, M. R. Genesereth, R. Fritzon, D. McKay, J. McGuire, P. Pelavin, S. Shapiro, C. Beck: Specification of the KQML Agent Communication Language, Technical Report EIT 92–04, Enterprise Integration Technologies, Palo Alto, USA, 1993.

[Gupta 94]
U. G. Gupta: How Case-based Reasoning Solves new Problems, *Interfaces*, Vol. 24, n.6, 1994.

[Jennings 93]
N. R. Jennings: Specification and Implementation of a Belief-Desire-Joint-Intention Architecture for Collaborative Solving, *Int. Journal of Intelligent and Cooperative Information Systems*, 1993, Vol. 2, N°3., pp 289–318.

[Pomerol et al. 93]
J. C. Pomero, B. Boy, C. Rosenthal-Sehoux and A. Saad: An intelligent DSS for the Multicriteria Evaluation of Railway Timetables, Euro Conference, Lisbon, Portugal, 1993.

[Schuster et al. 97]
A. Schuster, W. Dubitzky, K. Adamson, D. A. Bell and J. G. Hughes: Processing Similarity between a Mix of Crisply and Fussily Defined Case Properties, in *Applied Intelligence: Int. Journ of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, Vol.7, pp1–18, Kluwer Academic Publishers, US, 1997.

[Simon 77]
H. A. Simon H.A., *The new Science of Managerial Decision*, ed; Prentice-Hall, New-Jersey, 1977.

[Summons et al. 98]
P. Summons, G. Gibbon, H. Villanueva and J. Falconer: Causal Reasoning in a Medical Knowledge Based System, in *Proceedings of ECIS'98*, Aix en Provence, France, 1998, pp1154–1165.

[Van Bemmel et al. 97]
J. H. Van Bemmel, M. A. Musen, R. A. Miller and A.A.F. Van der Maas: Methods for Decision Support, chap 15. in *Medical Informatics*, ed. by J.H. Van Bemmel and M.A. Musen, Springer Verlag, 1997, pp. 233– 260.

[Wiederhold 92]
G. Wiederhold: Mediators in the Architecture of Future Information Systems, *IEEE Computer*, 1992, N°3, pp.38–49.

to Contents, page 1

Bioinformatics and Health: Impacts of the Human Genome Project on Health Informatics

Fernando Martín Sánchez

Medical Informatics, defined as the application of information technologies to the management and analysis of clinical information and bioinformatics, resulting from the union of Computer Science and Genetics, have been independent disciplines with methods, objectives and well differentiated curricula. However, with the advance the Genome Project and the association of our genes with the causes of the diseases, we are witnessing the integration of clinical information and genetics in different environments, from basic research to clinical practice and public health. In this work the convergence between Bioinformatics and Medical Informatics is described as also the technologies that will make it possible and the main applications of these new focuses and its potential impacts in health.

Keywords: Bioinformatics, Medical Informatics, Genomics, Biochips, Diagnosis, Pharmacogenomics, electronic clinical record.

Introduction

Biomedicine, an information-based discipline, is undergoing profound changes as the new experimental approaches generate enormous volumes of unprecedented data (Genome Project, clinical trials, medical images). Biology and medicine are increasingly looking for support from the application of the information sciences and technologies.

The huge amounts of information generated by the new genetic technologies are gradually released to the scientific community, mainly over Internet, and have to be integrated and analysed in order to extract biomedical knowledge useful for the development of new diagnostic and therapeutic solutions.

As the Human Genome is deciphered and the knowledge about diseases is associated with the genes involved in their development, it becomes increasingly clear that the integration of genetic and clinical information is going to change the face of medicine in the coming years.

This paper outlines the main issues related to genetic information processing (bioinformatics) and clinical information processing (medical informatics) and indicates how the two are converging into what could be termed biomedical informatics, analysing the main impacts of the Human Genome Project on the development of information systems for health professionals.

Fernando Martín Sánchez is Ph. D. in Computer Science and in Molecular Biologist. Postdoctoral at the Biomedical Informatics Program (Emory University Hospital, Georgia Tech, Atlanta, USA). Head of the Department of Bioinformatics and Public Health Informatics at the Institute of Health *Carlos III*. Assistant Professor of Bioinformatics at the University *Francisco de Vitoria* of Madrid. Member of the board of the Spanish Health Informatics Society (SEIS). His main research interests are in the development of computer systems to facilitate the new approaches of genomic-based medicine. <biotic@isciii.es>

The Human Genome Project

The discovery of the of the DNA (deoxyribonucleic acid) molecule structure by Watson and Crick in 1953 triggered a scientific revolution in biology and opened the door to what is now known as molecular genetics.

Information is stored in living beings as packets or quanta termed genes. A gene is a portion of genetic material that encodes the information required to create molecules that are to fulfil particular functions within cells. The information present in genes is used for protein synthesis. As a simile with computer sciences, DNA could be compared with the source code of computer programs, whereas the proteins would be the executable programs.

Genomics is the branch of biology pursuing the study of the genomes of organisms. A genome is defined as all the genetic information present in a given organism. The Human Genome

Figure 1: The Human Genome Project

to Contents, page 1

Project (HGP) is an international project that aims to build detailed physical and genetic maps of the genome of our species and identify the full sequence and location of the genes of which the genome is composed.

The scientific community is now setting new goals, which include using the huge amount of structural information generated so far in the development of functional analyses. So, what we are now witnessing is the transition from Genomic to the Post-Genomic Era, where the genomes and the relationships between their structure and their function will be analysed and compared.

The Human Genome Project (see Figure 1) is expected to be useful as a source of knowledge for understanding biological phenomena and diseases and lead to new methods of diagnosis, drugs and treatments for genetic-based diseases.

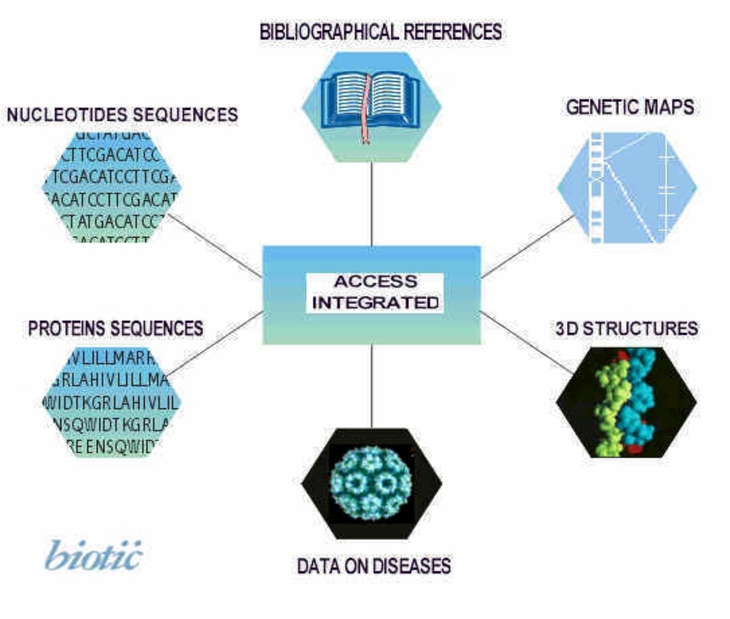


Figure 2: Integrated access to biomedical information

Bioinformatics

Bioinformatics is a relatively recent term and has not appeared in the literature before 1991 [Boguski 98]. Bioinformatics lies at the crossroads between the life and information sciences, providing the tools and resources required to promote biomedical research. It is an interdisciplinary scientific field, whose goal is to research and develop systems that are useful for giving an understanding of the flow of information from genes to molecular structures, their biochemical function, their physiological behaviour and, finally, their influence on diseases and health.

The main stimuli for the development of this discipline were the enormous volume of genetic data generated by all the genome projects (human and of other organisms), the new experimental biochip-based approaches, which output genetic data at high speed, and the development of Internet and the WWW, providing universal access to biological information databases.

The most prominent applications of work in this area are:

- Biological information databases
- Sequence alignment and comparison
- Protein structure and function prediction
- Genetic and metabolic maps
- Polymorphism and gene expression data analysis
- Phylogeny and evolution studies

The main challenge for bioinformatics is to provide a response to the avalanche of data from genomics. Whereas the results of experiments could be interpreted by means of a laboratory log only a few years ago, databases and data visualisation techniques are now needed to merely store and commence their analysis. Bioinformatics is evolving from just a series of techniques towards a full-blown science, as it provides the analysis component for understanding genomics and integrating genomic data (see Figure 2) to create predictive models of biological systems.

Biochips or DNA chips

Biochips are miniaturised devices with a high density integration of biological material located in an array, providing a means of rapid and efficient genetic analysis. They are called chips by analogy with the high density electronic circuits present in a microelectronic chip. They are having a big impact on research and have great clinical potential. They can be used to get individual genomic material and have the potential of providing a means of portable, rapid and economical diagnosis, which could be applied at the point of health care itself. [Martín 98].

Biochips are divided into small cells, each of which acts like a test tube where a reaction takes place. There are a huge number, as many as hundreds of thousands, of cells. Thanks to the extreme miniaturisation of the system, it is possible to analyse all the possibilities of mutation of a gene simultaneously on one chip or detect which genes are expressed at any time. The chip is put into contact with a fluorescent marked sample and is then washed, after which only the fragments of DNA that hybridise complementarily will remain linked. The chip is then

Figure 3: Images related to the use of DNA chips (Source: Affymetrix – BIOTIC)

entered into an optical scanner for the process of development, which will locate the strings marked with the fluorochrome (see Figure 3). A computer analyses the information from the scanner and outputs the result. (Wallace 97).

Individual genetic polymorphism

The sequence obtained thanks to the genome project will be a general sequence of the human species, as it corresponds to only the few people whose genetic material has been sequenced. However, the genetic make-up of different individuals is thought to be diverse, as a result of mutations and polymorphisms that amount to around 0.2% of our genome. It is precisely these differences that are very much related to people being more or less susceptible to developing diseases or to the varying therapeutic action of drugs.

Rapid sequencing systems can be used to examine all the possible polymorphisms and detect mutations in complex genes. The sequences of one and the same gene can vary from individual to individual, leading to a mutation. SNPs (single nucleotide polymorphisms) are caused by the mutation of a single DNA residue and can lead to a loss of gene function, a disease or to the acquisition of susceptibility to a disease.

The meaning of the variation of human genetics is analysed by observing the mutations of sequences of normal genes, which are then correlated with certain diseases or specific drug-response patterns.

Medical Informatics

Medical informatics is over 40 years old and has developed from a technology into a basic medical science. It includes the theoretical and practical issues related to the processing and communication of information derived from medical processes and related to health. It overlaps with almost all the medical specialities to configure a multidisciplinary sector with branches in epidemiology, technology evaluation, health economics and management and medical ethics.

Some of its main lines of development are:

- Computerised clinical records
- Education
- Hospital information and documentation systems
- Signal and medical image processing
- Knowledge-based systems for diagnosis and treatment support
- Surgery and radiation planning and simulation
- Health statistics and indicators
- Telemedicine
- Vocabulary, disease coding and procedures
- Evidence-based medicine support
- Epidemiology and public health system support

Emerging technologies

Interest of the big information and communications technology companies in bioinformatics

The major information and communications technology companies have decided to take a stake in the development of bioinformatic solutions for the future. They are building bioinformatic solutions centres or entering into cooperation with leading organisations or research centres in the field.

Compaq supplies Celera Genomics, which has completed the Human Genome Project sequencing, with its information technology infrastructures, allowing the company to use the Internet and electronic commerce to distribute its findings to the scientific community. Compaq has also supplied the Sanger Centre, a public institute located in Hinxton in the United Kingdom and which has sequenced 30% of the human genome and the genomes of a range of micro-organisms, with hardware.

IBM pursues its research activities in biological computing at its *T.J. Watson Research Center* in Yorktown Heights, New York. It has also entered into co-operation with the *Hospital for Sick Children* in Toronto, Canada, providing equipment for the *Genome Database* (GDB), one of the databases most extensively used by researchers from all over the world. IBM has announced an investment of $100 million earmarked for advanced sequence analyses and identifying proteins using the *Blue gene* supercomputer. IBM has also entered into agreements with bioinformatics companies like SBI (Structural Bioinfo) or Netgenics, which apply these new technologies to drug discovery.

SUN, again, has taken an interest in the intersection between biomedicine and informatics, launching the *Discovery Informatics Program* initiative, in which it is co-operating with a series of other companies (DoubleTwist, Curagen, GCG, Lion, Informax, Timelogic).

Silicon Graphics (SGI) has had strong links with the industrial and academic bioinformatics community for some time.

Motorola Inc. has signed an agreement with the Packard Instruments Company and Argonne National Laboratory to market and sell advanced biochips. Also Agilent (a subsidiary of Hewlett Packard) has announced an important initiative in the field of the health sciences: the genetic microarrays program (DNA arrays or biochips). Agilent markets a DNA and RNA-based bioanalyser Lab-Chip, developed jointly with Caliper. The bioinformatics platform that is to be used is based on the Rosetta Inpharmatics *Resolver* expression data analysis system.

The most promising computer technologies in this sector are described briefly below.

Data mining and visualisation

Together with new experimental approaches, the appearance of new technologies for outputting genetic information has led to the generation of huge quantities of data that must be managed and stored for later analysis using bioinformatics tools. This new challenge has led to the development of new methods, tools and to the application of data mining. Data mining involves extracting the predictive information that is latent in large databases. Several techniques and methodologies, including decision trees, clustering techniques, rule induction, neural networks and genetic algorithms, are used for this purpose. The models produced by this modelling or predictive work are visualised by means of special-purpose tools that employ a range of methods to ease their understanding.

Internet portals, e-genetics and e-health

Bioinformatics was one of the first scientific and technical disciplines to make use of the features of Internet-based systems to develop solutions and provide for data sharing, software distribution and co-operative work between groups and scientists [Brown, 2000].

With the explosive growth of the Internet and the completion of the first phase of the Human Genome project, there is a trend, running parallel to electronic commerce in other branches of activity, towards the appearance of Internet portals that give researchers accessibility to genetic data and bioinformatic tools. Craig Venter, leader of Celera Genomics, recently said that he manages an "Internet Company". Other big bioinformatics and genomics companies make their data and software, formerly reserved to their customers, available to universities, small and medium-sized companies and researchers through web interfaces, working as ASP (Application Service Provider), in some cases with commercial criteria. Additionally, these services are viewed by the large pharmaceuticals and biotechnological industries as the possibility of "outsourcing" bioinformatics so as not to have to make an enormous outlay in specialised technology and human resources [Fisher, 99].

Again, there are initiatives in the field of genetic diagnosis for the development of teleconsultation services by means of biochips over the Internet, or capturing patient genetic material for research purposes from medical Internet portals. Some are already predicting that in a few years people will be able to ask a laboratory for their genetic profile on a CD-ROM and compare this with web-based genetic libraries that will be able to inform them of their likelihood of suffering certain diseases or of the drugs they tolerate best.

Database integration, CORBA, components technology

The size of biological databases is growing at an exponential rate. The management of these resources and their use by scientists is becoming a far too complex task. Genetic databases are organised and structured very differently, although they often contain interrelated data. Some are supported by standard relational managers, while others are composed of plain-text files. One evident technological challenge is to integrate these sources of information, thus providing for their unified access irrespective of any changes of organisation they undergo.

CORBA and Java provide interoperability and portability in respect to Internet-based database access. Component-based technologies are valuable in that they can be used to extend bioinformatic applications across networks, operating systems and different programming languages, providing developers and users with a guarantee of scalability, maintenance and ease of use.

XML

There are over 600 databases of interest for genomics researchers. Many provide not only data but also services. Taking into account the differences there are in interfaces, syntax and semantics, it is almost impossible for a user to be able to use any databases of interest all at the same time. One of the biggest challenges of bioinformatics is to respond to the continuous growth and changes to these databases.

XML (eXtensible Markup Language) is a programming language that has emerged as an evolutionary successor to HTML and is put forward as the candidate technology for interconnecting all these sources of knowledge. It is used to describe the content of a document and has been designed as a SGML (Standard Generalized Markup Language) application, which makes it possible to create specialized documents. XML is suitable for any type of data description, which means that XML-based information systems can be used to create, store ad distribute information. An XML-based information system can contain different document type definitions (DTDs), which are formal document grammars, and users can, therefore, create documents that comply with particular DTDs.

Several XML-derived languages have appeared in bioinformatics: BIOML (BIOpolymer Markup Language) is a DTD for sequence and structure data in molecular biology, BSML (Bioinformatics Sequence Markup Language) is another DTD, which is used to represent molecular biology data, and MAML (Microarray Markup Language) for standardising gene expression data obtained with microarrays.

Knowledge management and ontologies

A conceptualisation is an abstract entity, which acts as a simplified view of the world that is to be represented for a purpose. Any knowledge-based system implicitly or explicitly includes a conceptualisation. An ontology is an explicit specification of a conceptualisation, that is, a description of the concepts and relationships that actually or possibly exist for an agent or a community of agents. Formally, an ontology is the establishment of a logical theory.

There is an important line of research for identifying technologies of use for managing knowledge and interchanging concepts and representations in biomedical sciences (Common Ontologies Exchange Language, genomic ontologies, drug research and development ontologies). This is of capital importance for organising the mass of knowledge generated daily on any subject of biomedical research.

Genomic impacts on clinical practice and health research

The application of genomic technologies (bioinformatics, biochips, etc.) and available biomedical knowledge is leading to the development of new approaches in the field of biomedical research, which are expected to be used in clinical practice in the coming years. Three of these application environments are outlined below.

Clinical diagnosis

The capabilities of the new tools used in genomic research are likely revolutionise clinical practice when applied as diagnostic tools. These high-performance new devices (biochips) enable a large number of parameters that can be used as diagnostic markers to be monitored simultaneously. This will have a big impact on the methods of genetic analysis. [Collins 99]. Around 6,000 genes associated with various diseases are

to Contents, page 1

known today and the list of genetically based diseases gets longer all the time.

Genetic analyses are used to identify a person is likely to suffer a disease, as well as to confirm a suspected mutation in an individual or a family. However, the growing interest raised by these field focuses on predictive analyses, which identify people with a high risk of contracting a disease before the associated symptoms appear.

Development of new drugs

Why do different patients with the same symptoms respond differently to the same drugs? Although the concept of individual variations on a molecular scale (pharmacogenetics) is not new (pharmacology has been addressing this problem from the toxicological viewpoint for 40 years), the technological platforms proposed for successfully implementing this concept have been developed in the last 4 years.

New technologies that ease the understanding of the role of genes in diseases is revolutionising the processes of new drug discovery and development (pharmacogenomics), providing the industry with substantial opportunities of saving time, cutting costs and lowering risks. The birth of personalised drugs for different strata of patients classed according to their genetic characteristics is near. The discovery of the genetic variants of individuals, which influence the effect of drugs, will enable the development of new diagnostic procedures and therapeutic products that will be selectively prescribed to patients with guaranteed safety and efficacy (Housman 98).

Genetic epidemiology

The use of new genetic information technologies will make cost-effective screening (genetic tests) possible at the population level. The challenge is to transfer the knowledge of genetics to the field of public health. For this purpose, it is important to assure the correct use of genetic information and new genetic information management technologies to produce health programs. There is talk of epidemiological/genetic study systems (associative genetics, genotype-phenotype population studies), and there are efforts aimed at disseminating genetic information, training health workers and developing policies that include the genetic knowledge output by the Human Genome Project in epidemiological practice.

The application of information technologies in this field can be helpful in carrying out cooperative efforts, like the HuGeNet (Human Genome Net) initiative, by the CDC's Office on Genetics and Disease Prevention [Khoury 98], to develop and distribute epidemiological information on the human genome, including:

– Specific prevalence data on genetic variants
– Epidemiological data on the relationship between genetic variability and diseases in different populations
– Evaluation of the validity and impact of genetic analyses.

Bioinformatics applications in health

Computer science has a fundamental role to play in the transfer of knowledge from genomics to health-care practice. The integration of genetic and clinical information, the training of health workers and patients and the support tools for managing all this new information is unthinkable without the use of information and communications technologies. The main field of application of bioinformatics, with its stock of genetic information processing and analysis, should be the clinical environment, although the most plausible hypothesis, as will be described later, is the convergence between medical informatics and bioinformatics.

Entry of genetic data into the computerised clinical records

The OMIM database contains over 6,000 genes associated with diseases. Genetic tests for several of these (Huntington's disease, cystic fibrosis, cancer) are already available. Although the analyses are now carried out in central laboratories and take several days, technologies can now transfer the genetic laboratory to the general practitioner's surgery, and even to the chemists'. This can be done thanks to the transition from PCR-based sequencing to new biochips, bringing us closer to the concept of an "automated clinical genotyping system".

With the advent and deployment of genetic information in routine clinical practice, associated tasks, like clinical records, will have to be adapted to be able to effectively manage this new type of information. The need for computer programs to assist in extracting, processing and managing these data will promote the inclusion of this information in clinical records. The benefits of including genetic information in clinical records are as follows:

- Application of medical knowledge to genetic sequencing for treatment
- Information sharing and transmission among practitioners
- Use in research – systems of associative genetics in epidemiology

Everything points to genetic information being routinely included in the computerised clinical records in a few years [Naser 98].

Genetic clinical practice guidelines

Another field in which the advent of genetic information is likely to cause a major impact is health management. Efforts will have to be made to develop new clinical practice guidelines and protocols, addressing genetic issues are included and to introduce new coding systems. Biochip-based technology can do for genetics what microprocessors did for computing. Thanks to the level of miniaturisation achieved, diagnostics will be able to move from the central laboratories into the practitioner's surgery, as computers left the computer centres and personal computers became ubiquitous. Clinicians are going to need regulated and negotiated procedures for developing the process of diagnosis and treatment. Computer science has been supplying excellent instruments for creating, disseminating and applying these clinical practice guidelines for years.

Continuous medical training

Biomedicine researchers and practitioners spend less and less time in the laboratory and library, while they increasingly use computers to store, interchange or analyse information. Even so, this convergence of biomedicine and computing is not reflected enough at universities and schools of medicine [Varmus 99]. If we acknowledge the quasi-ubiquity of the information communication and processing activities in medical practice, it is not difficult to appreciate how the progress made by these technologies can lead to qualitative changes in the way in which health care is provided. Therefore, the need to specially train practitioners in these subjects at undergraduate, postgraduate and continuous medical education has been recognised. No formal process of training practitioners in these subjects has yet been provided in this country, and everything seems to indicate that the information and communication technologies can provide a guarantee of continuous or on-the-job training, taking into account how quickly knowledge is updated in these fields.

Access of health workers to resources and genetic databases

Bearing in mind the enormous rate of genetic-based medical knowledge generation, it is not hard to imagine that physicians will be obliged to consult the Internet to locate up-to-date information required in routine practice (databases on genetics, diseases, scientific papers, clinical histories, epidemiological data, information on drugs, …) in a few years' time [Sikorski 97].

One of the major challenges for bioinformatics over the next century is to manage to link clinical information with molecular information. One especially important point is the need to manage information at health organisations that are going to undergo drastic changes owing to the developments in molecular genetics [Altman 98].

The data handled by bioinformatics will pervade in medical practice, and the hospital information systems will have to be prepared to house all this new range of data.

Redefining and coding diseases

Diseases were first grouped according to patient symptoms, a classification technique that is still used. Later, as knowledge progressed, the results of the analyses that provide a better and more accurate definition of the disease-causing agent were added to the clinical symptoms. Thanks to the knowledge supplied by genomics-based medicine, this system may well be redefined, and diseases will stop being defined on the basis of the phenotype (external traits, symptoms) and be identified by the causal mechanism and the patient genotype. On this point, diseases have already been identified that used to be considered different and which have been found to have the same origin (two dystrophies that were thought to be of different origin and that share the affected gene). Additionally, it will be discovered that different diseases can be interrelated (the apoE gene is involved in causing cardiovascular diseases and Alzheimer).

Computer-assisted diagnosis systems

Clinicians will be faced with new decision arrays on the best treatment, in which they will have to consider individual genetic variations and the molecular subtypes of the disease.

	Disease Subtypes					
Patient Genetic Traits	X					
						X
				X		
		X				

Table 1: Decision array for personalised treatment

Thanks to approaches like this, it will also be possible to better monitor the therapies applied, and ascertain traits that have an impact on the treatment response and efficacy of the drugs and which could sometimes suggest they be changed or omitted. The diversity of human genetics causes population groups with given characteristics that make them more likely to suffer from some diseases than others. This is precisely what will enable pharmaceuticals companies applying pharmacogenomic techniques to develop personalised drugs and medicines that will better suit certain molecularly characterised population groups. Computer science has to provide assistant or expert systems to help physicians with this complex task.

Conclusions

The convergence of medical informatics and bioinformatics. Opportunities for co-operation

Searching for the keyword *Informatics* in MEDLINE, the reference database of scientific papers published in biomedical journals, 346 publications appear in the last five months. Of these, 175, that is, more than half, are on genomics. This is indicative the growing importance of work on bioinformatics in health [Kohane 2000].

Medical informatics and bioinformatics are growing closer, as information on the human genome is generated and linked with medical knowledge on diseases [Altman 00]. This trend is evident from the fact that medical informatics congresses like MIE, AMIA or MEDINFO have special sessions on bioinformatics and the masters in medical informatics at several universities, like Stanford, include bioinformatics training.

Although bioinformatics and medical informatics share many methodologies, data types and algorithms, it is noteworthy that bioinformatics is more advanced in: [Altman 98]

- Database integration technologies
- Data representation systems and algorithms
- Systems validation methods and hypotheses

Medical informatics has developed further in areas like (Miller, 2000):

- Reasoning systems
- Terminologies, coding and vocabularies
- User interfaces
- Knowledge representation
- Data mining
- Modelling and simulation
- Probabilistic methods

There are a quite a few areas of common interest, where there synergy could take place [Kohane 2000]:

to Contents, page 1

- Data interchange format standardisation
- Components for distributed computing
- Image processing
- Organisation and access to scientific literature and sources of information
- Knowledge representation
- Data models
- Controlled vocabularies
- Processing of sources of data error and noise in signals
- Confidentiality and data security
- Interfaces for efficient data entry
- Knowledge discovery tools

Some examples of systems where the barrier between bioinformatics and clinical informatics is fuzzy are:

- Systems for processing individual genetic information and its role in disease development
- Integration of genetic information in clinical records and computer-assisted diagnostic decision-making systems
- Processing of data from biochips or DNA microarrays
- Pharmacogenetics databases [Altman 2000] (genetic information in clinical trials)

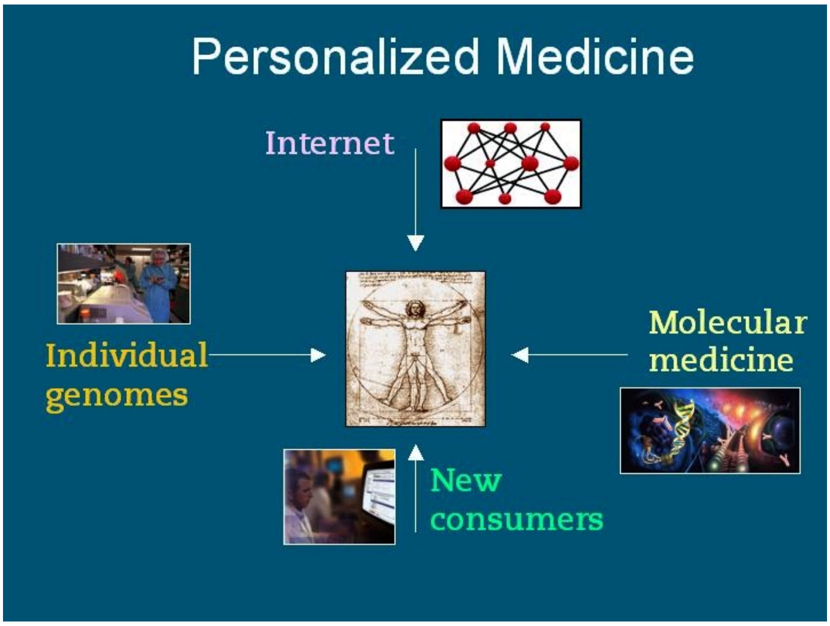


Figure 4: Factors supporting the development of personalised medicine

Personalised, preventive, molecular medicine

The main issues of what is likely to be a new revolution in medicine triggered by the introduction of genetic information and its integration into the clinical process have been presented. This could be characterised as:

- *Molecular* – We will witness the birth of what is known as "molecular medicine". Molecular medicine can be defined conceptually as the effort to define both normal and pathological physiological states in terms of the presence and regulation of the molecular entities making up living beings.
- *Preventive* – Prevention in this sense means the medical actions taken in respect of the environment and life style, aimed at reducing the risk of genetically predisposed individuals suffering from diseases. Knowledge of the genetic traits of populations would make it possible to ascertain the likelihood of suffering from certain diseases before the symptoms appear, thus enabling the implementation of improved and genuine preventive medicine.
- *Personalised* – Several forums are pointing out the importance of genomics-based approaches to furthering the personalisation of medicine. The recently completed sequencing of the human genome opens up enormous possibilities for developing new diagnostic procedures, treatments and drugs that will be selectively prescribed to patients, with guaranteed safety and efficacy.

Three recent prospective reports on medicine coincide in pointing out the impact of genomics and the Internet on future health systems. The papers "HEALTHCARE 2000: A strategic assessment of the Health Care Environment in the United States", "HEALTHCAST 2010" and "HEALTH and HEALTHCARE 2010, the Forecast, the Challenge" (see references), published almost simultaneously, give an understanding of the technological trends that are likely to have most impact on the health system in the coming years. The three reports stress especially the impact of genomics and biomedical informatics on the system of disease prevention, diagnosis and treatment. The development of personalised medicine (see Figure 4), adapted to the genetic particularities of patients and where diseases are classed according to their molecular cause, is one of the biggest challenges in the sector, and it is here that these techniques and methods are going to most useful. They will also contribute to the gradual transformation of palliative or curative medicine into genuine preventive medicine, where diseases can be detected and treated even before the first symptoms appear.

Ethical, legal and social issues

Despite their undeniable usefulness, studies of the genome involve some ethical risks that cannot be obviated and which must be carefully addressed. The private nature of genetic information raises serious obstacles to translating the knowledge of the human genome into benefits to the health of the individuals. In 1989, the very architects of the human genome project admitted that the information that would be obtained from sequencing the human genome and mapping genes would have profound implications for individuals, the family and society and established the ELSI program to anticipate and study the ethical, legal and social issues that would be raising as a result of research into human genetics. The Neufield Council was the main mover of the ethical debate in England. UNESCO has set up an International Committee of Bioethics and has promoted a Declaration on Genome-related Human Rights. The Council of Europe issued a report on the protection

to Contents, page 1

of human rights and the dignity of the human being with respect to biological and medical applications in 1997.

The migration of the new genetic technologies (microarrays) into clinical practice will make possible rapid and economic identification of people with a high risk of suffering a genetic disease in the near future. This is one of the most immediate challenges for the ELSI project. Before these technologies are applied, actions will have to be developed that assure the safety and quality of the tests and raise the knowledge of health workers concerning their benefits and risks. Additionally, patients must receive proper genetic advice by means of which to make a decision in full knowledge of its implications and individuals must be assured that the information will not be used against them or their families.

References

[Altman 98]
Altman, R.B. "Bioinformatics in support of molecular medicine". In C.G.Chute, Ed., *AMIA Annual Symposium*, Orlando, 1998. p. 53–61.

[Altman 00]
Altman, R. "The Interactions Between Clinical Informatics and Bioinformatics". JAMIA Vol 7 Number 5, 2000, p 439–443.

[Boguski 98]
Boguski, M.S.,."Bioinformatics-A new era". Trends guide to Bioinformatics,. 1, 1998: p.1–3.

[Brown 00]
Brown, S."Get your bioinformatics on the web". *Biotechniques*. 28, 2, 2000p244-246.

[Collins 99]
Collins F. S. Shattuck Lecture "Medical and Societal Consequences of the Human Genome Project". *The New England Journal of Medicine* 341, 1999:28–37.

[Deloitte/Touche 00]
Deloitte and Touche y HVA, 2000. "*HEALTHCARE 2000: A strategic assessment of the Health Care Environment in the United States"*, 2000

[Fisher 99]
Fisher L. "Surfing the Human Genome" *The New York Times on the Web* (Technology), 1999.

[Housman 98] Housman D. "Why pharmacogenomics? Why now?" *Nature Biotechnology* 16, 1998:492

[Institute for the Future 99]
Institute for the Future, "*The Future of the Internet in Health care*". Menlo Park. CA, 1999

[Institute for the Future 00]
Institute for the Future and Robert Wood Johnson Foundation, 2000 "HEALTH and HEALTHCARE 2010, the Forecast, the Challenge", 2000

[Khoury 98]
Khoury M.J. "HuGENet". 1st *Annual Conference on Genetics and Public Health*. 1998

[Kohane 00]
Kohane, I. "Bioinformatics and Clinical Informatics: The Imperative to Collaborate". *JAMIA* Vol 7 Number 5, 2000, p 439–443.

[Naser 98]
Naser C.R. "*High Speed Genetic Testing Technology and the Computerization of the Medical Record*", 1998.

[Martín/López-Campos 98]
Martín F. and López-Campos G. "Tecnologías basadas en Biochips. Aplicaciones en diagnóstico clínico e investigación biomédica". *II Symposium Internacional sobre Diagnóstico Genético en Medicina*. Madrid 1998.

[Miller 00]
Miller, P. "Opportunities at the Intersection of Bioinformatics and Health Informatics". *JAMIA* Vol 7 Number 5, 2000, p 431–438

[Price Waterhouse Coopers 00]
Price Waterhouse Coopers. "HEALTHCAST 2010", 2000

[Sikorski et al. 97]
Sikorski R. et al. "Genomic Medicine, internet resources for medical genetics". *JAMA* 278, 1997:1212–1213.

[Wallace 97]
Wallace R.W. 1997. "DNA on a chip: serving up the genome for diagnostics and research". *Molecular Medicine Today* 3, 1997: 384–389.

[Varmus 99]
Varmus, H. "Working Group on Biomedical Computing. The Biomedical Information Science and Technology Initiative". *National Institutes of Health* (U.S.A.), 1999.

[WHO 99]
World Health Organization. "*Human Genetics and noncommunicable diseases*", 1999.

to Contents, page 1

Secure Internet-Access to Medical Data

Ulrich Ultes-Nitsche and Stephanie Teufel

This paper reports on context-dependent access control to support distributed clinical trials. Centrally stored data will be accessed from contributors to a clinical trial over the Internet. We present in this paper how context-dependent access control can be implemented on the Internet in a secure way, using Java Servlets to implement the access control and SSL to secure communication.

Keywords: Workflow-support for clinical trials, telemedicine, context-dependent access, health information systems.

1 Introduction

In [Holbein et al. 97] and previous papers [Holbein/ Teufel 95], [Holbein et al. 96], the concept of a context-dependent access-control has been introduced and discussed exhaustively. In context-dependent access control, information about the state of a business process is combined with general knowledge about a user to grant or revoke access to sensitive data [Holbein 96]. A prototype implementation of the concept is described in [Nitsche et al. 98]. The prototype implementation is for local use only and would reveal many security holes if used over an open network. However, using technology different from the one presented in [Nitsche et al. 98] allows to come up with a secure distributed solution to context-dependent access control over the Internet.

In this paper, we present an implementation concept for context-dependent access control on the Internet. Even though applied to the specific application area of clinical trials, the underlying concepts are general and support all applications for which context-dependent access control is suitable. The paper summarizes a part of the Swiss National Science Foundation funded project MobiMed [Fischer et al. 95]. The system we are going to describe is PC-based (Windows NT) and implemented as a Java Servlet accessing an MS SQL Server database.

2 Context-Dependent Access Control

Role based security approaches fit well the hierarchically structured setting of a hospital [Ting et al. 91], [Ting et al. 92]. Each level in the hierarchy can be mapped to a so-called organizational role that a person at this hierarchical level plays in the hospital (e.g. medical staff, care staff, etc.). After an analysis of each role's demands on obtaining particular data to perform work, access rights are assigned to each role. The access rights determine which records in the database that contains information about the patients (patient records) may be read or written by a person playing a particular role. When logging in, a user of the system identifies her- or himself, ideally by using a chipcard and a PIN, and then a role is assigned to the person according to user/role lists, determining her or his access rights. A different way to handle role assignment is to store role information on the chipcard itself in ciphered form, which then is read during log in or, alternatively, whenever data records are accessed.

Simple role based access control mechanisms have the advantage that they can be implemented rather easily but the drawback of certain inflexibilities. A more sophisticated access control technique refining the role based approach takes into account an access request's particular point in time, i.e. the question: "Is it reasonable that a person playing a particular role needs access to certain data at the current state of a healthcare process?" Obviously, it is not necessary to have access to all data about a particular patient all the time. The question of "what does one need to know?" moreover, "what does one need to know right now?" (at the time of an access request) is the context-dependent access control scheme that we consider in this paper. A possible realization of the need-to-know principle

Dr. Ulrich Ultes-Nitsche is a lecturer at the University of Southampton, U.K. He studied Computer Science and Physics at the University of Frankfurt. Before joining Southampton, he was a junior scientist at GMD Darmstadt, a visiting researcher at the University of Liège, and a research assistant at the University of Zurich. In 1995 he received a DAAD HSP II/AUFE fellowship and in 1997 he received the Best Student Paper award of the 16th Annual ACM Symposium on Principles of Distributed Computing (PODC'97), Santa Barbara, USA. His research interests comprise security, formal methods, and automated verification. His web-page is at http://www.ecs.soton.ac.uk/ ~ uun.

Stephanie Teufel is professor in management in telecommunication and the director of the international institute of management in telecommunications (iimt) at the University of Fribourg. She studied informatics at the Technical University of Berlin and the ETH Zurich, where she received her diploma degree in Computer Science (MSc.). She has been a teaching fellow and lecturer at the University of Wollongong, Australia. Afterwards she was a senior researcher at the Department of Informatics of the University of Zurich, where she received her Doctor's degre and her Habilitation. She hold a professorship in business informatics at the Department of Informatics at the Carl von Ossietzky Universität Oldenburg. Major research interests: Mobile Electronic Business, Information Security Management, Information- and Communication Management, and Technology Management.
<stephanie.teufel@unifr.ch>

to Contents, page 1

can be achieved by combining user role information with state information of a workflow ("Is it reasonable to grant access to a person (playing a particular role) in the given workflow state?").

3 The Access-Control System

As mentioned above, context-dependent access control combines user information with process state information to compute access rights. Since the database system that we consider – MS SQL Server – uses group-based access control, we have to build context dependency around group-based access controls.

Access rights in the context-dependent scheme are only given temporarily to users for a single access. Therefore we create a not yet existing group, called here MobiMedDBAccessGranted. No user and no group is assigned to MobiMedDBAccessGranted *initially and permanently*. In order to change the group membership of a user temporarily, we use the SQL stored procedure sp_change_group. After checking the "need to know" of a user to perform an access [Holbein et al. 97], he or she receives temporary membership in group MobiMedDBAccessGranted, the access is performed, and the user is reset to her/his group.

As context information we use state information from a business process model. In such a model, a state represents the history of the process. To each patient a single instantiation of the business process is created. The process ID of the instantiation is used as a unique key to all database entries related to the patient. Context-dependent access-control tables are generated, containing quadruples (stateID, groupID, table_entry, access_type), indicating that in state stateID, a user belonging to group groupID can access database entry table_entry by an access of type access_type. Access types are e.g. *full access*, *read-only access* or *no access*. The presented access-control concept can be implemented on the Internet around Java Servlets and SSL.

4 Java Servlets

The past year has seen the rise of server-side Java applications, known as Java Servlets. Servlets are used to add increased functionality to Java-enabled servers, replacing CGI and offering many significant advantages [Sun Microsystems 98] (we focus on security here):

4.1 Portability

Java Servlets are protocol and platform independent and as such are highly portable across platforms and between servers.

4.2 Performance

Unlike CGI scripts, Servlets do not create a new process for each incoming request. Instead, Servlets are handled as separate threads within the server, reducing object creation overhead.

4.3 Security

The Java language and Java Servlets have improved security over traditional CGI scripts both at the language level and at the architecture level:

4.3.1 Language Safety

As a language Java is type safe and handles all data types in their native format. With CGI scripts most values are treated and handled as strings which can leave the system vulnerable. For example, by putting certain character sequences in a string and passing it to a Perl script, the interpreter can be tricked into executing arbitrary and malicious commands on the server.

Java has built-in bounds checking on data types such as arrays and strings. This prevents potential hackers from crashing the program, or even the server, by overfilling buffers, which is commonly known as stack smashing and can occur with CGI scripts.

Java has also eliminated pointers and has an automatic garbage collection mechanism, which reduces the problems associated with memory leaks and floating pointers. The absence of pointers removes the threat of attacks on the system where accesses and modifications are made to areas of server memory not belonging to the service process.

Finally, Java has a sophisticated exception handling mechanism, so unexpected data values will not cause the program to misbehave and crash the server. Instead an exception is generated which is handled and the program usually terminates neatly with a run time error [Garfinkel/Spafford 97].

4.3.2 Security Architecture

Java Servlets have been designed with Internet security issues in mind and mechanisms for controlling the environment in which the Servlet will run have been provided.

CGI scripts generally have fairly free access to the server's resources and badly written scripts can compromise the security of a server by either leaking information about the host system that can be used in an attack, or by executing commands using untrusted or unchecked user arguments. Java significantly reduces these problems by providing a mechanism to restrict and monitor Servlet activity. This is known as the Servlet sandbox. The Servlet sandbox provides a controlled environment, in which the Servlet runs, using a security manager to monitor Servlet activity and prevent unauthorized operations.

In JDK 1.2 an extension to its security manager, the access controller, is introduced. The idea behind the access controller is to allow more fine-grained control over the resources a Servlet can access. For example, instead of allowing a Servlet to have write permission to all files in the system, write permission can be granted for only the files required by the Servlet for execution [Hunter 98].

However, Java-based servers are still vulnerable to denial of service attacks where the system is bombarded with requests in order to overload the server resources. However, the effects of this can be reduced by specifying an upper limit on the number of threads that can be run concurrently on the server. If all the threads are allocated, that particular service can no longer be accessed, but because the server still has resources left to allocate, the rest of the services are still available.

5 The Secure Sockets Layer Protocol

The secure sockets layer protocol (SSL) is designed to establish transport layer security with respect to the TCP/IP protocol stack. Version 3 was published as an Internet draft document [Freier et al. 96] by the IETF (Internet Engineering Task Force). In combination with Java Servlets, the use of SSL is studied in [Ultes-Nitsche 00].

Unlike other concepts that secure connections or even only data-packages, SSL includes the concept of a secure session, determined by the parameters negotiated at the beginning of the session by the two protocol machines. It is this session concept that makes it appealing for being used in the MobiMed prototype. The secure session lasts as long as a user is logged in the system. Since communication with the user will be based on HTML documents sent to and received from the client side using http, the use of SSL will be transparent to the user.

6 Implementing the Access-Control System

In order to perform the SQL stored procedure sp_change_group, the Servlet is equipped with SQL Server administrator privileges. The Servlet offers the only way to access the MobiMedDB database and can be accessed from any application on the Internet. By putting the SQL Server administrator information into the private part of the Servlet class, it is securely protected.

Java Servlets support the Java DataBase Connectivity (JDBC) API to access databases that support the JDBC API, too. SQL Server supports the Open DataBase Connectivity (ODBC) API that can be linked to the JDBC API. The interfaces within the resulting system are presented in Figure 1.

The client-side application talks to the Servlet via the webserver. The connection between client and web server is secured using SSL. The Servlet API handles http get and post requests. From the SQL Server it receives user information as well as the access-control tables for the evaluation of access rights. The Servlet handles user log-ins as well as setting and resetting group membership. An example of a successful request under the context-dependent access control scheme is described in eight steps in Figure 2 (AW Administrator is the runtime environment of the workflow system):

1. The user (application) sends an access request to the Servlet.
2. The Servlet requests and receives context information about the workflow of the patient whose record is attempted to be accessed.

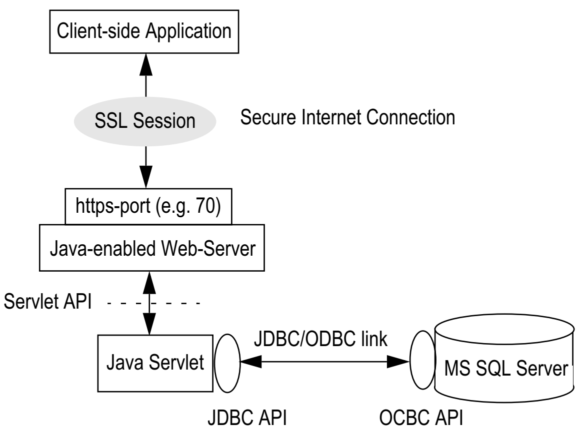


Fig. 1: The system inferfaces.

3. The Servlet requests and receives group information about the user who is requesting the access.
4. Both, context and role information are combined and the authority of the access is evaluated. Subsequently it is assumed that the user has authority to perform the request (otherwise the request would be rejected at this point).
5. The user is set temporarily to group MobiMedDBAccess-Granted.
6. The request (query) is performed on the database MobiMed-DB.
7. The results of the query are delivered to the user.
8. The Servlet resets the user to his/her previous group membership.

7 Discussion

To ensure privacy, the SQL Server and the Java-enabled Web server, including the access-control Servlet, must form a single system in which only the Web server is accessible remotely. Data exchange between client and server is confidential by using SSL. SSL also supports authentication of client and server.

Given that the access-control system is implemented as described, the obvious major threat to the system is that the administrator log-in procedure known to the Servlet becomes publicly accessible. By hosting the sensitive parts of the access control in a private method of the Servlet and taking into account that Servlets are server-side Java bytecode, neither can a user access the private Servlet methods directly nor is bytecode containing sensitive information accessible from the Internet. The sensitive parts of the system are hence securely kept.

A way a user could try to break the system is by sending multiple access requests, knowing that at least for some of them she/he has authority. The idea is that a request for which she/he has no authority coincides sufficiently in time with a request for which she/he has authority. So both queries could be send to the SQL Server at the time when she/he is temporarily assigned to the group for which access is granted. This situation can only occur, when multiple threads of the access-control Servlet are not synchronized properly. The simplest attempt to solve this problem is to disable multi-threading for single users.

Since accesses to the system are relatively rare compared to heavily used Internet servers, performance is not very much an issue in our case. Tests with an early local prototype [Nitsche et al. 98] and a similar system [Röhrig/Knorr 00] did not show any problems with performance.

It should be noted that in the setting of a hospital, access to data can be vital. Therefore context-dependent access control will be equipped with simple, group-based bypass mechanisms for emergencies. However, bypassing the context-dependent access control will have to be logged thoroughly to provide proof of potential misuse of the bypass mechanism.

8 Conclusion

In this paper we reported on a project on context-dependent access control to support distributed clinical trials. We concentrated on presenting the technical aspects of the solution, in

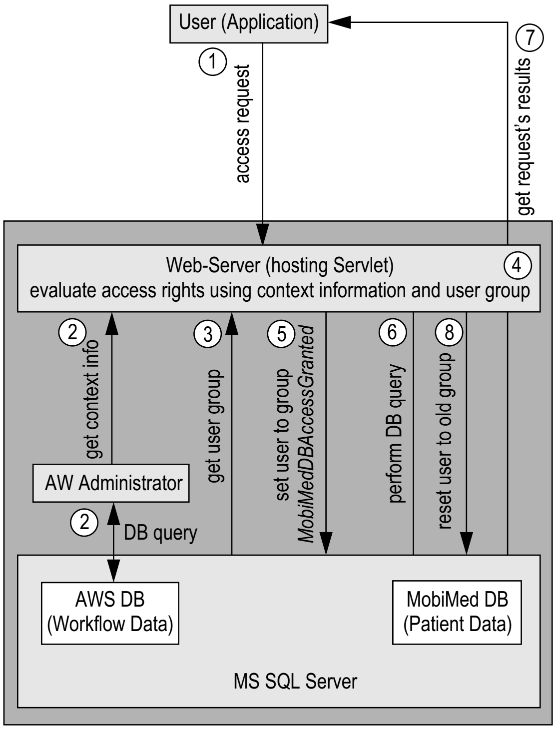


Fig. 2: The events comprising a successful access request.

particular on the use of Java Servlets. The implementation concept comprises a secure distributed solution to context-dependent access control [Nitsche et al. 98], [Röhrig/Knorr 00] over the Internet. The proposed system is PC-based (Windows NT/ Windows 2000), using a Java Servlet to access an MS SQL Server database (that contains the medical data) in a context-dependent fashion. To secure the communication the secure sockets layer protocol (SSL) is used. The underlying security concepts presented here are general and in particular the Internet security issues can be adapted to different platforms [Ultes-Nitsche 00] and applications [Hepworth/Ultes-Nitsche 99] without change.

Acknowledgements

We thank Prof. Dr. Kurt Bauknecht and the SNSF for their support and Konstantin Knorr for helpful comments on an earlier version of this paper.

References

[Fischer et al. 95]
H.-R. Fischer, S. Teufel, C. Muggli, and M. Bichsel. *Privacy and Efficiency of Mobile Medical Systems (MobiMed)*. Bewilligtes Forschungsgesuch des SNSF Schwerpunktprogramms Informatikforschung SPP-IuK, Modul: Demonstrator, Nr. 5003-045359, 1995.

[Freier et al. 96]
A. O. Freier, P. Karlton, and P. C. Kocher. The SSL protocol version 3.0. In Internet Draft http://home.netscape.com/eng/ssl3/ draft302.txt, Netscape, Transport Layer Security Working Group, November 1996.

[Garfinkel/Spafford 97]
S. Garfinkel and G. Spafford. *Web Security and Commerce*. O'Reilly and Associates, 1997.

[Hepworth/Ultes-Nitsche 99]
E. Hepworth and U. Ultes-Nitsche. Security aspects of a Java-Servlet-based Web-hosted e-mail system. In *Proceedings of the 7th IFIP Working Conference on Information Security Management & Small Systems Security, 1999*. Kluwer Academic Publishers, Boston.

[Holbein 96]
R. Holbein. *Secure Information Exchange in Organizations*. PhD thesis, University of Zurich, Switzerland, 1996. Shaker Verlag, Aachen.

[Holbein/Teufel 95]
R. Holbein and S. Teufel. A context authentication service for role based access control in distributed systems – CARDS. In J. Eloff and S. von Solms, editors, *Information Security – the Next Decade IFIP/SEC'95*. Chapman & Hall, London, UK, 1995.

[Holbein et al. 96]
R. Holbein, S. Teufel, and K. Bauknecht. The use of business process models for security design in organisations. In S. Katsikas and D. Gritzalis, editors, *Information Systems Security-Facing the Information Society of the 21st Century (IFIP/SEC'96)*. Chapman & Hall, London, UK, 1996.

[Holbein et al. 97]
R. Holbein, S. Teufel, O. Morger, and K. Bauknecht. A comprehensive need-to-know access control system and its application for medical information systems. In L. Yngström and J. Carlsen, editors, *Information Security in Research and Business (IFIP/SEC'97)*. Chapman & Hall, London. UK, 1997.

[Hunter 98]
J. Hunter. *Java Servlet Programming*. O'Reilly and Associates, 1998.

[Nitsche et al. 98]
U. Nitsche, R. Holbein, O. Morger, and S. Teufel. Realization of a context-dependent access control mechanism on a commercial platform. In G. Papp and R. Posch, editors, *Global IT Security – Proceedings of the 14th International Information Security Conference (IFIP/Sec'98)*, volume 116 of OCG, Vienna, Austria, 1998.

[Röhrig/Knorr 00]
S. Röhrig and K. Knorr. Towards a secure web-based health care application. In H.R. Hansen, M. Bichler, and H. Mahrer, editors, *8th European Conference on Information Systems (ECIS'2000)*, Vienna, Austria, 2000.

[Sun Microsystems 98]
Sun Microsystems. *Java Servlet API Whitepaper.* 1998.

[Ting et al. 91]
T. Ting, S. Demurjian, and M.-Y. Hu. Requirements, capabilities, and functionalities of user-role based security for an object-oriented design model. In C. Landwehr, editor, IFIP WG 11.3 *Workshop on Database Security*, Sheperdstown, West Virginia, 1991. Elsevier Science Publishers.

[Ting et al. 92]
T. Ting, S. Demurjian, and M.-Y. Hu. A specification methodology for user-role based security in an object-oriented design model. In B. Thuraisingham and C. Landwehr, editors, *IFIP WG 11.3 6th Working Conference on Database Security,* Simon Fraser University Burnaby, Vancouver, British Columbia, 1992. Elsevier Science Publishers.

[Ultes-Nitsche 00]
U. Ultes-Nitsche. Secure server-side Java applications. In *Informatik/Informatique – Journal of the Swiss information processing societies* 6 / 2000. Special Issue on Security and the Net.

to Contents, page 1

Data Warehouse Dissemination Strategies for Community Health Assessments

Donald J. Berndt, Alan R. Hevner, James Studnicki

The Comprehensive Assessment for Tracking Community Health (CATCH) methodology provides a systematic framework for community-level assessment that can be a valuable tool for resource allocation and health care policy formulation. CATCH utilizes health status indicators from multiple data sources, using an innovative comparative framework and weighted evaluation process to produce a rank-ordered list of critical community health care challenges. The community-level focus is intended to empower local decision-makers and provide a clear methodology for organizing and interpreting relevant health care data.

Keywords: Data Warehousing, OLAP (On-Line Analytical Processing), Health Care Informatics, Decision Support Systems, Database Systems.

1 Community Health Organizations

It is well documented that considerable variation exists in the health status of defined populations. This variation is evident when we compare large population groups, such as separate nations, states, or regions within a single country. Surprisingly, variation often persists within smaller population groups, such as census tracts or zip codes inside United States counties. These variations exist not only for what would be considered epidemiological health status outcomes (i.e., morbidity and mortality rates), but also for indicators which could be considered other dimensions or domains of population health such as socio-economic and demographic characteristics, the availability of health resources, patterns of health behaviours, and many other factors. In order to improve the health status of populations, a continuous monitoring and improvement system must be implemented. Such a system requires a comprehensive, objective, and uniform methodology for defining and characterizing the many dimensions that comprise the health status of a community.

As part of the on-going clarification of the public health role at the community level and the transitions from a *disease* to a *health* focus and from a *treatment* to a *prevention* strategy, there has been recognition that partnerships and collaboration are necessary to support effective action [Institute of Medicine 96], [Nakajima 97]. Health organizations, public sector agencies, medical care providers, businesses, the religious community, educational institutions, and other community organizations are interdependent components of a multi-sectoral community health organization. The overall community must be empowered to make the necessary, and sometimes difficult, resource allocation choices to improve health through information, education, behaviour change, and social support [Cropper 96]. Such collaborative action at the community level must be informed by *unbiased data* describing the community's health

Donald J. Berndt is an Assistant Professor in the Information Systems and Decision Sciences Department in the College of Business Administration at the University of South Florida. He received his M.Phil. and Ph.D. in Information Systems from the Stern School of Business at New York University. He also holds a M.S. in Computer Science from the State University of New York at Stony Brook and a B.S. in Zoology from the University of Rhode Island. Dr. Berndt's research and teaching interests include the intersection of artificial intelligence and database systems, knowledge discovery and data mining, data warehousing, software engineering methods, and parallel programming. <dberndt@coba.usf.edu>

Alan R. Hevner is an Eminent Scholar and Professor in the Information Systems and Decision Sciences Department in the College of Business Administration at the University of South Florida. He holds the Salomon Brothers/Hidden River Corporate Park Chair of Distributed Technology. His areas of research interest include information systems development, software engineering, distributed database systems, health care information systems, and telecommunications. He has published over 80 research papers on these topics and has consulted for a number of Fortune 500 companies. Dr. Hevner received a Ph.D. in Computer Science from Purdue University. <ahevner@coba.usf.edu>

James Studnicki is a Professor in the College of Public Health and Director of the Centre for Health Outcomes Research at the University of South Florida. He received a B.S. in Psychology from the University of Pittsburgh, an MBA from George Washington University, and a Master of Public Health and a Sc.D. from Johns Hopkins University. His areas of research interest include community health needs assessment, outcomes associated with alternative medical treatments, medical practice guideline development, dissemination and impact, cybernetic systems for professional behaviour change, gender and race/ethnicity associated with treatment patterns, and predictors and correlates of medical malpractice.<jstudnic@hsc.usf.edu>

* * *

This research is partially funded by the U.S. Dept. of Commerce through a Technology Opportunities Program (TOP) grant.

to Contents, page 1

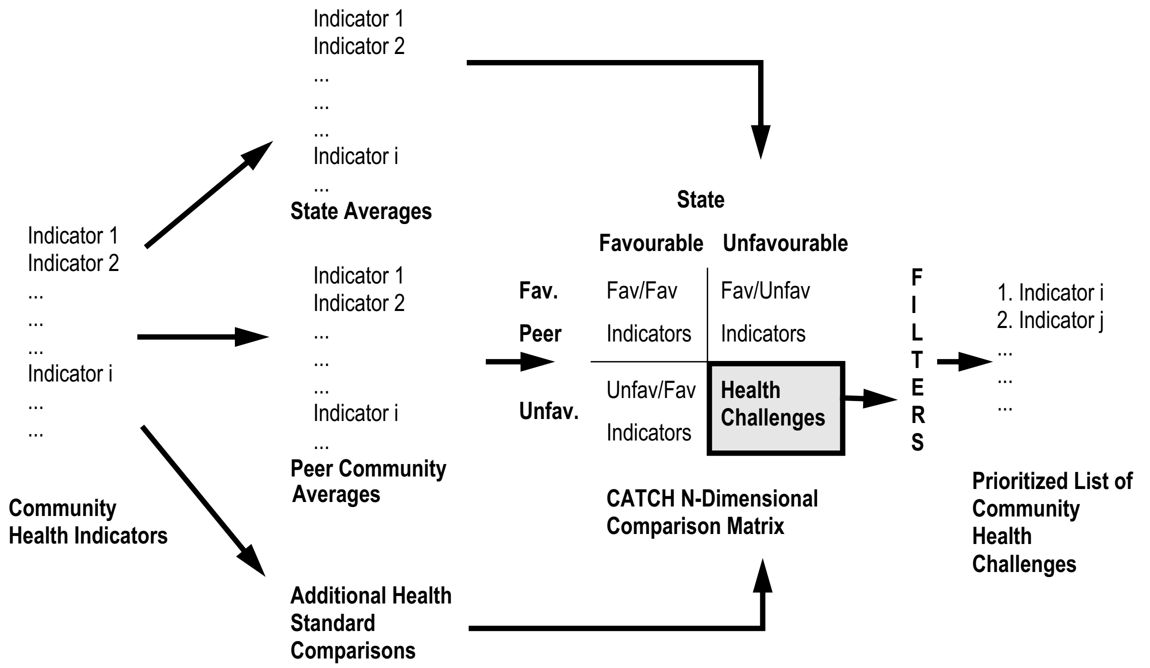


Fig. 1: The CATCH Methodology

status, needs, and resources. The ability is also needed to track progress over time to meet the community's health care goals.

The gap between current practice in community health care spending and the above goals of collaborative community health care decision-making is vast. The availability and quality of data on health indicators are problematic. There is little empirical evidence on the use, sharing, or strategies to integrate health data into decision-making to provide guidance to community health organizations. While most of the literature on collaborative leadership and community engagement focus on the process [Centre of Disease Control and Prevention 97], [Chrislip/Larson 94], little attention has been focused on the effect of the availability of a common set of data, such as the community health profile, on the quality and inclusiveness of decision-making. There is also scant information about the use of data and information technology to support and monitor the process.

The purpose of this paper is to present an outline of the Comprehensive Assessment for Tracking Community Health (CATCH) methodology and its implementation in a data warehouse. The various modes of data dissemination from the data warehouse to the community are explored and examples of current CATCH interfaces are demonstrated. We conclude by examining important issues of community decision support on health care.

2 The CATCH Methodology

The University of South Florida's Centre for Health Outcomes Research has developed the CATCH methodology to provide comprehensive, objective health status data for community health planning purposes. CATCH collects, organizes, analyses, prioritizes, and reports data on 225 health and social indicators on a local community basis. The CATCH methodology has been tested, refined, and validated over the past nine years. Reports have been prepared for 15 U.S. counties both within and outside of Florida.

The CATCH methodology can be briefly described as shown in Figure 1. Community health indicator data are gathered from a variety of sources. Secondary data sources include health care data reported by hospitals, local, state, and federal health agencies, and national health care groups. Primary data sources would involve data gathered from door-to-door or mail-in surveys. All health care data are normalized into common formats and organized into a community health care report card listing values for each important community indicator.

Each indicator value is then compared against the state average, a peer group of communities' average, and other interesting values (e.g., a national goal for that indicator). The results of these comparisons are organized into an n-dimensional matrix based on favourable or unfavourable comparisons against each comparison dimension. Figure 1 shows a 2-dimensional comparison matrix based on state averages and peer averages. Community indicators that demonstrate unfavourable comparisons on all dimensions are highlighted as community health challenges. This set of health challenges are prioritized by passing each indicator through a set of ranking filters:

- Number Affected – Number of persons in the community affected by the indicator.
- Economic Impact – An estimate of the direct cost per case for individuals affected by the indicator.
- Availability of Efficacious Intervention – An estimate of the relative degree to which treatment or prevention is likely to be effective.

to Contents, page 1

- Magnitude of Difference – The degree to which the community indicator is worse than the dimensional comparisons.
- Trend Analysis – For a five-year period is the trend favourable or unfavourable and what is the magnitude of change in the trend direction?

The community stakeholders are given an opportunity to weight the importance of each of the above filters. The final product of the CATCH methodology is a comprehensive, prioritized listing of community health care challenges. A more detailed description of the CATCH methodology with a complete listing of health care indicators can be found in [Studnicki et al. 97].

3 CATCH Data Warehouse

3.1 Limitations of Manual CATCH

While the value of CATCH is incontrovertible, the ultimate deployment of CATCH throughout Florida and the nation has been constrained by serious limitations.

- The hand crafted process is labour-intensive and slow. Hundreds of individual sources of data must be identified, a variety of formats reconciled, and data quality checked by hand. With current methods, it takes 3 to 4 months to complete a CATCH report for a single county.
- Longitudinal trend analyses over many years are cost prohibitive for most communities since each application is expensive and time-consuming.
- Most public health funding comes from state and federal governments. A statewide CATCH assessment would help to prioritize funding and serve to enable effective program evaluation based on quantifiable outcomes assessment. However, an automated methodology is critical for widespread adoption.

3.2 CATCH Data Warehouse Challenges

A CATCH data warehouse has been constructed to overcome these limitations, enabling both cost-effective report generation and ad-hoc analyses of critical health care issues. The construction of a data warehouse for public health care data poses major challenges beyond that required for the construction of a commercial data warehouse (e.g., retail sales).

- Data come from a diverse set of sources. Health care data are published in a wide variety of formats with differing semantics and there are few standards in the health care field for data.
- CATCH reports are disseminated to a diverse and geographically distributed set of stakeholders.
- The data warehouse is required to support the activities of public policy formulation. The sociopolitical issues of health care policy impact design features such as security, availability, data quality, and performance.

3.3 Data Warehouse Design

Important missions of a data warehouse include the support of decision-making activities and the creation of an infrastructure for ad-hoc exploration of very large collections of data. Decision makers should be able to pursue many of their investigations using browsing tools, without relying on database programmers to construct queries. The emphasis on end-user data access places a premium on an understandable database design that provides an intuitive basis for navigating through the data. The star schema or dimensional model has been recognized as an effective structure for organizing many data warehouse components [Gray/Watson 98]. The star schema is characterized by a centre fact table, which contains numeric information that can be used in summary reports. Radiating from the fact table are dimension tables that provide a rich query environment. This structure provides a logical data cube, with dimensions such as time and location identifying a set of numeric measurements within the cube.

The most appropriate facts are *additive* numeric data items that can be summed, averaged, or combined in other ways to form summary statistics. The only way to "compress" the millions of transaction items is to present some mathematical summarization. No human will want thousands, let alone millions, of items in answer to their queries. As Kimball points out, "The best and most useful facts are *numeric*, *continuously valued*, and *additive*" [Kimball 96].

The mission of the CATCH data warehouse is to support the automated and cost-effective application of the CATCH methodology, as well as to enable more detailed analyses that were not possible using the coarse-grained data that typified past CATCH reports. In order to meet these goals, the data warehouse design includes several levels of data granularity, from the coarse-grained data used in generic report production to actual event-level data, such as hospital discharges. The data warehouse includes major components at three levels of granularity, as illustrated by the data access pyramid in Figure 2.

1. Reporting tables with highly aggregated data are used to support the core CATCH reports, including comparisons between the target county and peer counties. These tables also provide fast response for interactive access via data browsing tools and can provide the foundation for simple community-wide Internet access.
2. There are families of star schemas that provide true dimensional data warehouse capabilities, such as interactive roll-up and drill-down operations. These components have care-

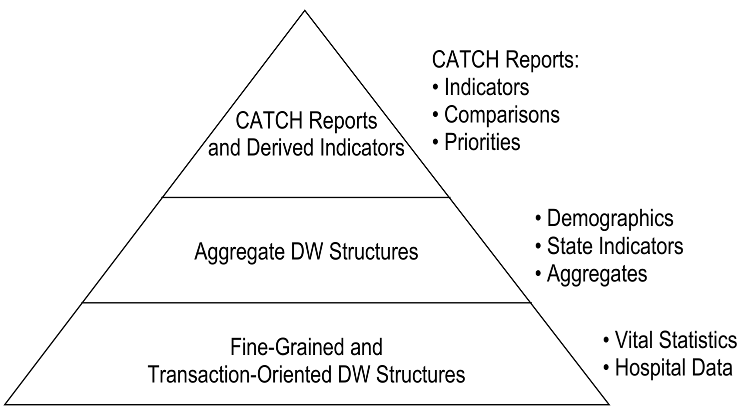


Fig. 2: The Data Access Pyramid

to Contents, page 1

fully designed dimensions that can be utilized by more sophisticated data browsing tools. The star schemas are populated using thorough data staging and quality procedures that usually involve processing detailed data sets extracted by various health care agencies and organizations. Typically, the data is aggregated and transformed for loading into a family of related star schemas that share important dimensions and support interactive online analytic processing (OLAP) techniques.

3. For certain types of information, the design calls for retaining very fine-grained or even event-level data. An example is the hospital discharge data that includes each hospital discharge event for the more than 200 hospitals that are mandated to report such information in Florida. This data is retained at the transaction level because of the rich set of facts and dimensions available for analysis and the density of potential aggregations that result in negligible space savings.

These three levels of aggregation within the data warehouse combine to meet a wide range of reporting requirements and performance goals. Thus providing a flexible basis for disseminating health care information to community decision-makers.

ing the results in tabular or graphical form. Though online analytic processing (OLAP) usually incorporates roll-up/drill-down features, the navigation style is highly interactive and driven by previous steps in data exploration.

Summarization tasks are algorithmic in nature, applying techniques that summarize patterns in the data and usually produce models, often with some notion of reliability, which can be used to predict as well as describe the underlying data. Traditional statistics and data mining techniques are often used as summarization tools. A distinction is drawn between their uses as *exploratory* or *confirmatory* methods, but the results are a model or set of abstract patterns that can be applied to other data sets. For example, connectivity to statistical packages is an important interface component that allows analysts to use statistical techniques to confirm or more fully investigate interesting properties discovered through browsing in the CATCH data warehouse. While these techniques are clearly important and applicable to health care data warehousing, the following discussion focuses on the navigation tools and more traditional database access technologies being utilized in the project.

4 Modes of CATCH Data Dissemination

The human-computer interface is of paramount importance in the data warehouse environment and the primary determinant of success from the end-user perspective. In order to support analysis and reporting tasks, the data warehouse must have high quality data and make that data accessible through intuitive interface technologies. The act of releasing data in a warehouse is in a very real sense the same as publishing that data in printed form; retractions in both media can be very painful. Once the data becomes accessible, it may be included in reports, forecasts, and analyses that form the basis of decision-making activities within an organization or community. Therefore, data staging and quality procedures within the data warehouse are often among the most expensive and critical ingredients in providing a successful end-user experience.

The types of access in a data warehouse can be broadly categorized as either *navigation* or *summarization* tasks. Navigation activities include data browsing, ad-hoc queries, and traditional report generation. These tasks require human guidance and design to produce the appropriate queries, often present-

USF University of South Florida Health Sciences Center

Indicator Fact Sheet

Indicator : First Trimester Prenatal Care - Total

County : Escambia

Peers: Alachua, Bay, Putnam

Indicator Category : Maternal and Child Health

Numerator Source: Florida Vital Statistics

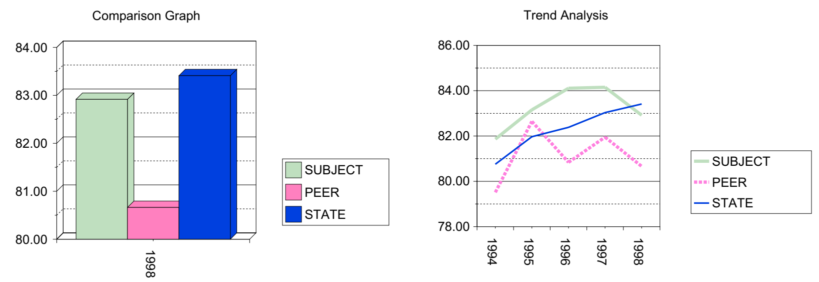
Denominator Source: Florida Vital Statistics

Description : Births to all mothers who indicated they received prenatal care (PNC) during their first three months of pregnancy

Calculation: Total live births with first trimester prenatal care/total live births x 100

Indicator Rate				Indicator Raw Values		
Current Year	Subject County	Peer Average	Florida	Subject County	Peer Average	Florida
1998	82.92	80.66	83.41	3,218	1,420	158,969

% of total live births

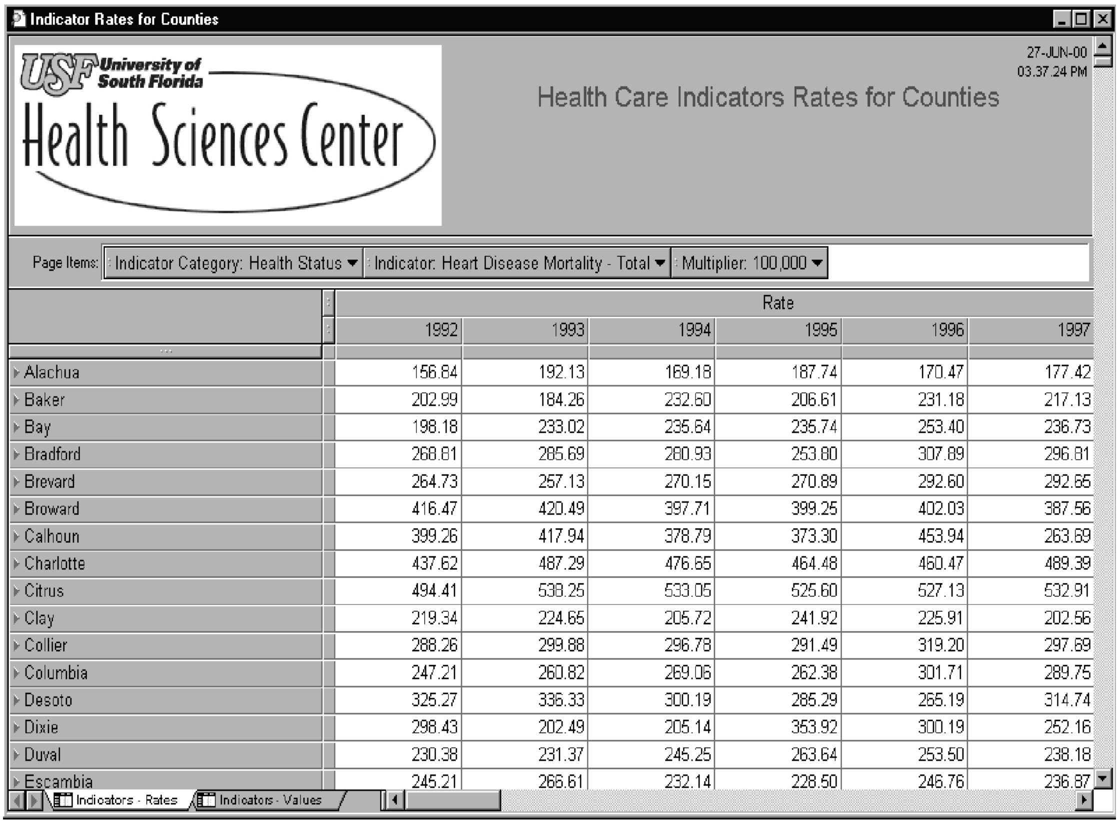


Indicator Rate				Indicator Raw Values		
Year	Subject County	Peer Average	Florida	Subject County	Peer Average	Florida
1994	81.86	79.51	80.75	3,263	1,406	153,111
1995	83.15	82.65	81.97	3,184	1,483	154,325
1996	84.11	80.84	82.38	3,239	1,441	155,747
1997	84.15	81.95	83.03	3,340	1,471	159,175
1998	82.92	80.66	83.41	3,218	1,420	158,969

% of total live births

Healthy People 2010 : 90% receive care in 1st trimester

Fig. 3: An Example Indicator Fact Sheet


Indicator Rates for Counties

27-JUN-00
03.37.24 PM

University of South Florida
Health Sciences Center

Health Care Indicators Rates for Counties

Page Items: Indicator Category: Health Status | Indicator: Heart Disease Mortality - Total | Multiplier: 100,000

	Rate					
	1992	1993	1994	1995	1996	1997
Alachua	156.84	192.13	169.18	187.74	170.47	177.42
Baker	202.99	184.26	232.60	206.61	231.18	217.13
Bay	198.18	233.02	235.64	235.74	253.40	236.73
Bradford	268.81	285.69	280.93	253.80	307.89	296.81
Brevard	264.73	257.13	270.15	270.89	292.60	292.65
Broward	416.47	420.49	397.71	399.25	402.03	387.56
Calhoun	399.26	417.94	378.79	373.30	453.94	263.69
Charlotte	437.62	487.29	476.65	464.48	460.47	489.39
Citrus	494.41	538.25	533.05	525.60	527.13	532.91
Clay	219.34	224.65	205.72	241.92	225.91	202.56
Collier	288.26	299.88	296.78	291.49	319.20	297.69
Columbia	247.21	260.82	269.06	262.38	301.71	289.75
Desoto	325.27	336.33	300.19	285.29	265.19	314.74
Dixie	298.43	202.49	205.14	353.92	300.19	252.16
Duval	230.38	231.37	245.25	263.64	253.50	238.18
Escambia	245.21	266.61	232.14	228.50	246.76	236.87

Indicators - Rates | Indicators - Values

Fig. 4: A Browsing Interface for County Level Indicators

4.1 Reports and Derived Indicators

The data warehouse components at the top of the pyramid provide the high-level reporting capabilities that were the focus of the original hand crafted assessments. The CATCH methodology has been refined through extensive field experience and provided a strong guiding framework for the initial data warehouse design. These top-level components include hundreds of indicators that are derived from the fine-grained data at the base of the pyramid. The identification of relevant indicators, describing the required calculations from underlying data, and understanding the use of information in the field are all valuable assets that have accrued through experience. This experience has been captured in many ways. The layout of the reports, OLAP interfaces, and the written analyses have all been refined over time. Most importantly, the hundreds of calculations necessary to derive the high-level indicators are implemented as stored procedures in the data warehouse and can be easily applied to new data or combined for novel approaches.

The derived indicators are used in the comprehensive community health assessment reports. Reports allow quick and easy access to comprehensive summaries and more detailed collections of information from the data warehouse. Figure 3 shows a fact sheet for one health indicator, with hundreds of such sheets being generated for each report. This type of pre-defined and thorough reporting is critical for implementing a more automated CATCH methodology. For example, the comparisons of target counties to peer counties and to the state are fundamental components of the original CATCH reports and important tools for community health care planners. The CATCH methodology has traditionally been centred on a large hardcopy report; so much of this content can be re-created in Web-friendly form and easily disseminated to local health planners. The advantage in this approach is the continued role of a strong methodology, rather than simply distributing raw data with no guidance in how to apply analytic methods.

4.2 Data Browsing and Aggregate Structures

Data warehouse browsing tools provide star query-like access through a flexible menu-based interface, with pull-down menus representing important data dimensions. These types of tools are easy to use and support ad-hoc exploration, but are usually controlled through some sort of administrative layer that determines the data available to end-users. In developing a flexible interface, there is a tradeoff between the ability to express ad-hoc queries and the ease-of-use that results from pre-defined constructs implemented by data warehouse designers and administrators. In addition, most of these tools are Web-enabled, providing dynamic access to a wide community through the Internet.

As noted in the data warehouse design discussion, the CATCH data warehouse consists of several levels of granularity from transaction-oriented data, such as hospital discharges, to summary data at the CATCH report level. Browsing tools can be used at all levels of the pyramid, but interface requirements will differ for each of the major components. For

to Contents, page 1

instance, the browsing tools provide a convenient method for CATCH analysts to view the preliminary report results with more detailed information than most community planners would want to sift through. Figure 4 shows an indicator over time for selected counties. Final report components may be generated using browsing tools, or more likely be implemented as part of a reporting function that more fully automates the process as discussed above.

A second and in some ways more important role for the browsing tools is to provide a flexible interface for more customized analysis. Aggregate data warehouse structures can often be used to improve browsing performance by physically instantiating the results of common roll-up operations. Custom analysis may also require additional calculations and aggregate structures are often an ideal location for such supplemental information. Example issues that are being investigated in this manner include racial disparities in infant mortality and the impact of surgical volumes on quality of care.

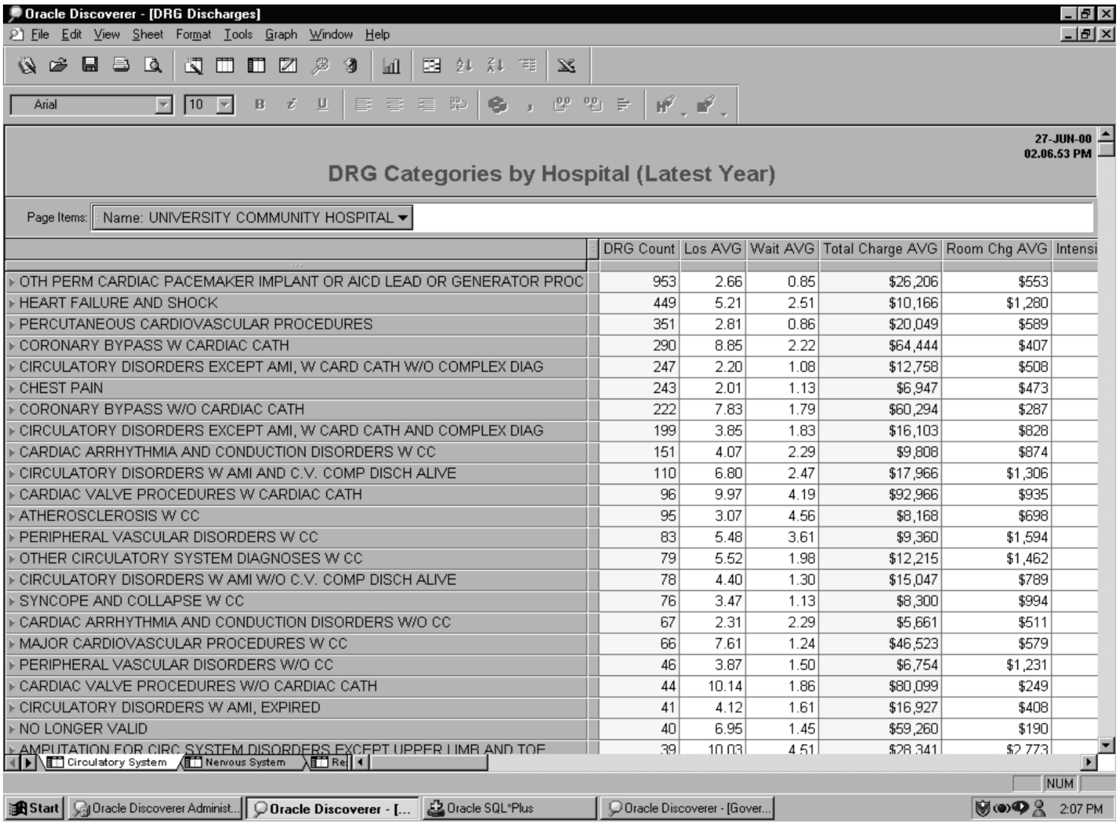
4.3 Fine-Grained Data Warehouse Structures

Health care issues highlighted by the CATCH methodology can be investigated more fully using the finer levels of detail maintained in the data warehouse. These tasks often entail querying the true dimensional star schemas that include age, gender, race, and other dimensions, or even the event-oriented data, such as hospital discharges. Thus the data warehouse allows the user to focus on issues such as differences in age or race with regard to specific health status indicators. Once decision-makers review the CATCH report, they may have specific issues that relate to the diverse communities that inevitably fall outside of arbitrary political boundaries. Figure 5 illustrates a detailed browsing screen in which volume, length of stay, and cost data are presented for a specific hospital by disease categories using diagnostic related groups (DRGs). It is clear how a hospital could effectively use this data for in-depth analyses of utilization and management decision-making.

5 Community Decision Making with CATCH Data

The CATCH data warehouse will result in widespread distribution of data previously unavailable to most communities as well as on-line access for specialized inquiry. This paper has focused on the rich variety of dissemination strategies available for making health care data available to local communities. Once received, however, many issues arise as to how the communities will make most effective use of the CATCH data for health care decision-making. This is an area with considerable research potential.

There is a rich literature on the decision-making process both with and without information technology. The study of group decision support systems and environments has a strong tradition in the management information systems field [Dennis 96]. In many ways, this important body of work is appropriate to health care decision-making that is usually group oriented. For example, the research in [Dennis et al. 98] studies the effects of minority influence on decision-making and finds that the presence or absence of technology has very different effects.



	DRG Count	Los AVG	Wait AVG	Total Charge AVG	Room Chg AVG	Intensi
OTH PERM CARDIAC PACEMAKER IMPLANT OR AICD LEAD OR GENERATOR PROC	953	2.66	0.85	$26,206	$553	
HEART FAILURE AND SHOCK	449	5.21	2.51	$10,166	$1,280	
PERCUTANEOUS CARDIOVASCULAR PROCEDURES	351	2.81	0.86	$20,049	$589	
CORONARY BYPASS W CARDIAC CATH	290	8.85	2.22	$64,444	$407	
CIRCULATORY DISORDERS EXCEPT AMI, W CARD CATH W/O COMPLEX DIAG	247	2.20	1.08	$12,758	$508	
CHEST PAIN	243	2.01	1.13	$6,947	$473	
CORONARY BYPASS W/O CARDIAC CATH	222	7.83	1.79	$60,294	$287	
CIRCULATORY DISORDERS EXCEPT AMI, W CARD CATH AND COMPLEX DIAG	199	3.85	1.83	$16,103	$828	
CARDIAC ARRHYTHMIA AND CONDUCTION DISORDERS W CC	151	4.07	2.29	$9,808	$874	
CIRCULATORY DISORDERS W AMI AND C.V. COMP DISCH ALIVE	110	6.80	2.47	$17,966	$1,306	
CARDIAC VALVE PROCEDURES W CARDIAC CATH	96	9.97	4.19	$92,966	$935	
ATHEROSCLEROSIS W CC	95	3.07	4.56	$8,168	$698	
PERIPHERAL VASCULAR DISORDERS W CC	83	5.48	3.61	$9,360	$1,594	
OTHER CIRCULATORY SYSTEM DIAGNOSES W CC	79	5.52	1.98	$12,215	$1,462	
CIRCULATORY DISORDERS W AMI W/O C.V. COMP DISCH ALIVE	78	4.40	1.30	$15,047	$789	
SYNCOPE AND COLLAPSE W CC	76	3.47	1.13	$8,300	$994	
CARDIAC ARRHYTHMIA AND CONDUCTION DISORDERS W/O CC	67	2.31	2.29	$5,661	$511	
MAJOR CARDIOVASCULAR PROCEDURES W CC	66	7.61	1.24	$46,523	$579	
PERIPHERAL VASCULAR DISORDERS W/O CC	46	3.87	1.50	$6,754	$1,231	
CARDIAC VALVE PROCEDURES W/O CARDIAC CATH	44	10.14	1.86	$80,099	$249	
CIRCULATORY DISORDERS W AMI, EXPIRED	41	4.12	1.61	$16,927	$408	
NO LONGER VALID	40	6.95	1.45	$59,260	$190	
AMPUTATION FOR CIRC SYSTEM DISORDERS EXCEPT UPPER LIMB AND TOE	39	10.03	4.51	$28,341	$2,773	

Fig. 5: A Browsing Interface for Hospital Information

Another important contributing area would be the political process and its ramifications to decision-making [Mintzberg 73]. Certainly, policy making in health care is very much a political process.

The use of the CATCH methodology and state-of-the-art data warehousing technology across many Florida communities will provide a rich research opportunity for studying many interesting issues on group decision-making in community health care organizations. Some of the issues we plan to study include the presence of a champion, group composition, socio-political context, and the ease of access and usefulness of the data. The complexities of each issue and the interrelationships among these issues make the design of research studies both a challenge and an opportunity.

Acknowledgments

The Florida Dept. of Health is a research partner. Research collaborators in the CATCH project include E. Gilbert, R. Marsh, D. McCorkel, B. Myers, M. Nevrekar, M. Pearl, J. Slayton, B. Steverson, and S. Hedge-Desai.

References

[Centre of Disease Control and Prevention 97]
Centre for Disease Control and Prevention, *Principles of Community Engagement*, Atlanta, 1997.

[Chrislip/Larson 94]
D. Chrislip and C. Larson, *Collaborative Leadership: How Citizens and Civic Leaders Can Make a Difference*, San Francisco, Jossey-Bass, 1994.

[Cropper 96]
S. Cropper, "Collaborative Working and the Issue of Sustainability," In: C. Huxham, editor, *Creating Collaborative Advantage*, London: SAGE Pub., 1996.

[Dennis 96]
A. Dennis, "Information Exchange and Use in Group Decision Making: You Can Lead a Group to Information But You Can't Make It Think," *MIS Quarterly*, 20(4), 1996.

[Dennis et al. 98] A. Dennis, K. Hilmer, and N. Taylor, "Information Exchange and Use in GSS and Verbal Group Decision Making," *Journal of MIS*, 14(3), Winter 1998.

[Gray/Watson 98] P. Gray and H. Watson, *Decision Support in the Data Warehouse*, Prentice Hall, Englewood Cliffs, New Jersey, 1998.

[Institute of Medicine 96]
Institute of Medicine, "Healthy Communities: New Partnerships for the Future of Public Health," Washington DC, National Academy Press, 1996.

[Kimball 96]
R. Kimball, *The Data Warehouse Toolkit*, John Wiley & Sons, New York, 1996.

[Mintzberg 73]
H. Mintzberg, *The Nature of Managerial Work*, New York, Harper and Row, 1973.

[Nakajima 97] H. Nakajima, "Editorial: New Players for a New Era," *World Health*, 50(3): 3, 1997.

[Studnicki et al. 97] J. Studnicki, B. Steverson, B. Myers, A. Hevner, and D. Berndt, "Comprehensive Assessment for Tracking Community Health (CATCH)," *Best Practices and Benchmarking in Healthcare*, September/October: 196–207, 1997.

to Contents, page 1

Software Visualization: An Overview

Luis M. Gómez Henríquez

This article provides a general overview of software visualization. To that end, we first present the concept of visualization and its features, and later we summarize some of its main applications in Computer Science. The main part of this article is devoted to describing the general discipline of software visualization, by defining it and its main problems and uses. Throughout the paper we will refer to some of the most outstanding actual systems which we think are suitable to illustrate the corresponding concept.

Keywords: Visualization, Software Visualization, Program Debugging, Performance Debugging

1 Introduction

Software Visualization (SV) is a discipline as ancient as Computer Science. However, it emerged as a independent CS field in about the mid 80s. In such early times, SV systems were mainly designed mainly to visualize the dynamic behaviour of algorithms. Its scope was widened to deal with other aspects of the programs, like static and dynamic views of their source code, or, as in another example, quantitative information about their parallel execution.

The objective of this paper is mainly to give an introduction, providing an overall view of SV discipline. In order to meet this goal, we try to generalize the most significant aspects of real SV systems instead of merely describing their particular features. This article contains a brief summary of our knowledge on the matter, which began with our experience in the design and development of Vestal [Tomás et al. 91] for Ada concurrent programs and VisAndOr [Carro et al. 93] which visualizes parallel execution of Prolog programs. Both systems and some theoretical work converged on our Ph. D. Thesis [Gómez 95]. In 1998, the first book devoted to SV was published: *Software Visualization: Programming as Multimedia Experience* [Stasko et al. 98]. However, this book, in our humble opinion, is no more than a compilation of papers about specific SV tools. Only the brief section introductions have some "systematic" content.

On the other hand, we have decided to include some basic information about the field of visualization and its possible applications. This matter has developed a lot in recent years. This fact and the improvement in graphic means, both software and hardware, have drawn on the emergence of promising new disciplines like Scientific Visualization and, more recently, Information Visualization.

2 Visualization

Both *to visualize* and *visualization* are usually heard in daily life. However, many times these words are improperly used: "Yesterday I was 'visualizing' *[sic*] the video I borrowed from you". Shortly said, to *visualize* is simply to give some visual aspect to an entity which lacks it. Consequently it is quite incorrect to use *visualize* as synonymous of *seeing* or *observing*: the result of *visualizing* is a *visualization* which can be seen or observed. Well, this *visualization* enjoys some virtues we are going to briefly discuss.

In general, it is said visualization allows us to make evident the *meaning* of an abstract entity, like a mathematical function or a huge amount of numerical data from a physical experiment or from a system simulation. The goal is make evident the significant occurrences, i.e., the "main trends" which could be difficult to discover by other means.

In recent years, what we can name *visual processing* of information has been a matter of active researching and discussion. *Perception Psychology* is the discipline which studies this subject.

Since the mid 60s, from the publication in 1966 of the work by the Nobel Prize winner (1981) Roger W. Sperry titled *Hemisphere Disconnection and Unity in Conscious Awareness*, it is thought the human brain has two different processing systems which are placed on both hemispheres. Though there is no common agreement on this placement, most scientists think the human brain has both a visual processing system (in short, in accordance with Sperry, "Right Brain") and a verbal one ("Left Brain"). Figure 1 (from the bestseller *Drawing on the Right Side of the Brain* by Betty Edwards) summarizes both systems characteristics.

In general, our system of communication, the human language, involves an intensive use of the "Left Brain". Consequently the "Right Brain" has been becoming less used from our childhood. Visualizations are clearly an object of the visual processing system or "Right Brain". Its characteristics are the same which can be seen in reality: 1) visualizations can significantly reduce the "seek time" for a given information (the seek

Luis M. Gómez Henríquez is an Associate Professor at the Dept. *Architectura y Technología de Sistemas Informáticos* of the *Universidad Politécnica de Madrid* where he received his Ph D in Computer Science. All of his research work has been related to the application of graphics in different aspects of Computer Science. His main current research interest is Information Visualization. <lgomez@fi.upm.es>

operation will be in "parallel"); and 2) a huge amount of information (which can even occupy many pages in its traditional textual form) can appear in a single picture.

About the first question, *preattentive processing* is produced by reducing this seek time to the lowest. Roughly speaking, preattentive processing is discovering some graphical fact within a time lower than that to begin to activate our rational attention system. This time has been estimated to range between 200 and 250 ms.

To conclude this section on visualization, we want to cite some interesting general books, though they are not directly CS-related ones. Briefly, about "mental organizers", [Buzan 96]; about the visual display of statistical data, [Cleveland 93] and the periodical publication *Statistical Computing & Statistical Graphics*; and at last "the Tuftes" [Tufte 83], [Tufte 90], [Tufte 97], three wonderful and fascinating books.

Left	Right
Verbal	Non verbal
Analytic	Synthetic
Symbolic	Concrete
Temporal	Non temporal
Rational	Non ratioal
Sequental	Parallel
...	...

Fig. 1: "Left brain" and "Right brain"

3 Visualization in Computer Science

Computer Science has always made use of visual representation techniques for different goals. However, only the progress in computer graphics allowed the emergence of "visualization" as an independent field in Computer Science. Examples range from the *pretty-printing*[1] of programs [Baecker/Marcus 90] to visual specification techniques in Software Engineering.

For example, *Scientific Visualization* has developed a lot since the mid 80s. Many works and publications about the subject can be found (see for example [Keller/Keller 93]). From a practical point of view, Scientific Visualization is usually supported by specific software. Some examples are: Data Explorer (DX) from IBM and AVS, both of them commercial, and Geomview and VTK (free distribution software).

Another previously cited discipline is *Information Visualization*. This discipline appeared in the 90s. The objective now is to visualize any kind of information, not necessarily a numerical one. The goal is usually to discover some occurrence in a short time. (IEEE CS organises an annual conference, *InfoVis*, since 1995).

Recently, the Data Mining world has made use of visualization techniques as well. These visual interfaces facilitate by specifying the requested information. Examples are the commercial systems *Clementine*, *Intelligent Miner*, and *MineSet*.

To finish, it is interesting to mention the intensive use of visual representations in Software Engineering. All Software Engineering models, from traditional *Booch diagrams* to *Rational Rose* or *UML*, use some kind of graphical pictures to represent the software system and its components.

1. A *pretty printer* is a program that formats a high-level program source text according to indentation rules based on the syntax and visualizing the program's structure.

4 Software Visualization

There is not a commonly accepted name for this discipline. In the past, by reviewing the existing literature, we found more frequently the terms *Algorithm Animation* and *Program Visualization*. However Software Visualization is currently the most usual denomination.

Neither is there a formal definition of it. In the reference book [Stasko et al. 98] we find: *"Software Visualization is the use of the crafts of typography, graphic design, animation, and cinematography with modern human-computer interaction and computer graphics technology to facilitate both the human understanding and effective use of computer software"*. In [Gómez 95] we proposed a simpler definition: "*Program Visualization is the art of giving programs another aspect than that of their source code*".

This visualization can be made from many different points of view: static (the source code) vs. dynamic; the algorithm of the program (its semantic) vs. the (non-semantic) evolution of its data structures; etc. On the other hand, Visual Programming is a related area: the visual specification of a program is also a visualization of its source code.

A new taxonomy of SV appears in [Stasko et al. 98]. This classification separates algorithm animation from other visualizations. We have always claimed that both algorithm visualization and those which show what happens at the hardware level are only different abstraction levels for observing the execution. A more detailed classification can be found in [Gómez 95]. Figure 2 shows a summary of it.

Regarding systems which show the execution of programs, this can be observed from different levels and perspectives. We distinguish three main levels: algorithm, source code, and hardware. In some levels it is possible to distinguish some sub-levels as well. Also different perspectives can be distinguished: for example, in the source code level we can observe the execution from the point of view of the data structures evolution (a recent example is *GNU DDD*) or, on the other hand, the control flow.

Some common characteristics can be found for the different levels and perspectives. These are: standardization degree (which ranges from lowest at the algorithm level to highest at

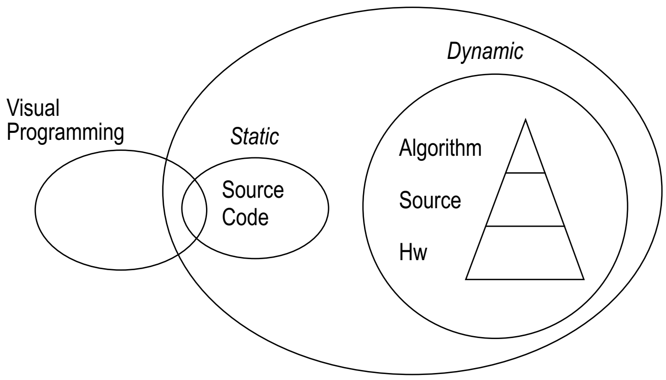


Fig. 2: Classification Software Visualization

hardware level), the use of the visualization (see section 7), and what we have named *temporal behaviour*: what procedure is used to represent time: animation, time-line, small multiples, etc [Gómez 95].

To conclude this introductory section to SV, we must be honest by saying SV has traditionally suffered from a lack of scientific recognition. In general, it has been said most of systems are "toys" without actual application in the real world. However, this opinion is changing. The arrival of some quality systems like the previously mentioned *GNU DDD* or other robust tools like *Imagix 4D* or *Javainsight* is achieving a radical change in such opinions.

5 A history of SV

In 1963 Donald Knuth published an article in *Communications of the ACM* about the computer generation of flow-charts. This is probably the first published work on SV.

In the early 70s some knowledge on how to visually display the way-of-working of programs appeared. The most famous one is [Baecker 75]. This work was developed by Ronald Baecker and it is probably the first one to show (through a film) the behaviour of an algorithm. Anyway, this visualization is fully hand-made and *ad hoc*.

In 1984 the first system capable of visualizing generic algorithms appeared. *Balsa* [Brown/Sedgewick 85] allows us to visualize the execution of programs written in Pascal. The source code is modified by the visualization designer by inserting animation primitives provided by *Balsa*. A similar system is *Tango* [Stasko 90]. This system has the additional features of allowing smooth movement of graphical objects.

The first system to visualize execution of concurrent or parallel programs appeared in the late 80s. *Voyeur* and *Belvedere* are two significant examples. A bit later *Zeus*, successor of *Balsa*, and *Polka*, successor of *Tango*, and *Pavane* [Catalin-Roman/Cox 89], the first system that employs a declarative "instrumentation" of programs, appeared. All of them visualize execution of concurrent programs.

ParaGraph [Heath/Etheridge 91] was constructed in 1991. It is still a "must-reference" tool. This system visualizes parallel execution from different perspectives, most of them show quantitative information. *ParaGraph* was the beginning of a very large set of similar performance tuning tools. This is probably the level (hardware) where the largest number of successful tools have appeared.

In 1994 some experiments on the use of Virtual Reality for parallel performance debugging was done at the University of Illinois. However we have seldom found any further work in this direction.

At the same time, a few experiences on the "auralization" or "sonification" of programs have been published. The main idea is to get program operations during execution converted to sound changes. In [Stasko et al. 98, chap. 10] you can find some references about this issue.

It is very difficult to summarize the large number of SV systems that have appeared since this time. We can say it is interesting to realise that some commercially robust tools are appearing. The previously mentioned *Imagix 4D* and *Jinsight* are two excellent examples. Another out-standing system is *PV* [PV] from IBM, mainly intended for performance debugging tasks. Another significant new system (for algorithm animation) is *Leonardo* [Leonardo], developed at University of Rome.

6 Some difficulties in SV

Visualizations of programs, and more specifically of their executions, suffer from some difficulties which it is necessary to mention when talking about SV. We briefly present some of them:

1) Program instrumentation

In order to visualize program execution it is necessary to gather significant execution events. The term *instrumentation* is usually used to name the process of adding the necessary mechanisms for gathering this execution information. Ideally this instrumentation must not modify execution behaviour at all. However, this goal is very difficult to achieve. The situation is quite similar to the "probe effect" in experimental sciences (someone called it "Heisenberg's Uncertainty principle applied to software"). This situation is even more complicated in the case of parallel programs. In such programs, a very small modification in their execution can significantly affect the overall behaviour.

2) Visualization definition

Particularly in algorithm visualizations, where the visualization paradigm can be as ad hoc as wanted, specifying the visualization can be very difficult. This specification is usually done by means of a set of graphical objects and their corresponding transformations, provided by the visualization system. These objects are created and transformed by inserting primitives in the source code – quite an arduous task. There are a few experiences on getting a visual specification of visualizations. John Stasko [Stasko et al. 98, chap. 14] describes some experiences in this sense. However, in our opinion this remains as an open question.

3) The scalability problem

This term is used to designate the capacity of adaptation of one entity to the change in the cardinality of some of its elements. This is one of the most challenging questions in the general world of visualizations. Specifically this has been a very studied question in the world of Information Visualization. Some partial solutions like semantic zoom or fish-eye views have been proposed.

4) The "missing-link" problem

In SV there are only two "touchable" realities: the source code and the resulting visualization. In consequence, establishing a "cause-effect" relation between both entities will be particularly useful. We call this relation the "missing-link". Only some recent systems like *Zstep95* [Stasko et al. 98, chap. 19] and the already mentioned *Imagix 4D*, both situated at the source code level, provide access the corresponding source code from the visualization level. It is interesting to note this problem is particularly difficult in both algorithm and hardware levels, where visualization present elements can be difficult to identify in program source code.

7 Use of SV

We mentioned above the fact that SV has traditionally been considered a "second-class" discipline. Opinions are usually found which say most systems are only academic experiments without any practical application. We must partially agree with these critics. Particularly in the algorithm level, the examples provided are only small programs which implement some classical algorithm, like sorting or bin-packing. However, it is difficult to imagine how to visualize a program from the real world, a lot more complex and where the algorithm concept can be very "diffuse".

This situation is less severe when the standardization degree is higher. In the case of the source code, nowadays we can find some commercial systems of high quality and usefulness, for example *Imagix 4D* or *Jinsight*. The situation is even better in the case of hardware-level visualizations. As we said above, this is the level where there are more successful tools.

Generally speaking, we can say SV systems are mainly used for the following purposes: 1) Program behaviour exhibition, normally for pedagogical purposes; 2) Logical debugging; and 3) Performance debugging.

The first one is usually created in the algorithm level. There are some observations in order to evaluate the degree of appropriateness of the use of the visualization system: "the learning time fell down from T to 2/3 T" or "80% students preferred to learn by using the visualization system", etc.

With regard to "logical" debugging, the traditional one (said to be the second oldest task in Computer Science), until recently only some experiments and discussion could be found, mainly for parallel programs. We think one of the main questions to solve here (despite the mentioned instrumentation one) is the need for guarantee that the graphical output is really what is happening in the actual execution. We think only some new systems like *Jinsight* are really reliable for this task. For some years the ACM/ONR Workshop on Parallel and Distributed Debugging was held. Some papers about the application of visual techniques can be found in the corresponding proceedings.

The third use is what we have named "performance" debugging. This task is particularly important in parallel programming. It is the main (and successful!) use of hardware-level tools.

8 The future

It is probable that SV will reach its definitive consolidation in the next years. One of the reasons we are so optimistic is the radical change that has happened since the mid 90s. SV has left its academic origin to become a discipline of commercial interest. There are still some open questions: the direct specification of visualizations, how to reduce the "probe effect", how to guarantee that the generation of visualizations is reliable enough to guarantee what you see is what is really happening, etc. New horizons are also appearing. The possibility of the "direct-manipulation" of visualizations, to allow the user to modify from the visualization some element of the program or its input (*Zstep95* is a first interesting experience in this sense). Virtual Reality will have a higher presence in SV: now we can imagine climbing an array column or seeing (and listening to!) the pass of messages in MPI.

References

[Baecker 75]
R. Baecker. Two Systems Which Produce Animated Representations of the Execution of Computer Programs. ACM SIGCSE Bull. 7. 1975

[Baecker/Marcus 90]
R. Baecker, A. Marcus. Human Factors and Typography for More Readable Programs. ACM Press. 1990.

[Brown/Sedgewick 85]
M. Brown, R. Sedgewick. Techniques for Algorithm Animation. IEEE Software 2(1). 1985.

[Buzan 96]
T. Buzan. Mind Map Book. Plume/Penguin. 1996.

[Carro et al. 93]
M. Carro, L. Gómez, M. Hermenegildo. Some Paradigms for Visualizing Parallel Execution of Logic Programs. Proc. of the Tenth Int. Conf. On Logic Programming, ICLP '93. The MIT Press. 1993.

[Catalin-Roman/Cox 89]
G. Catalin-Roman, K. Cox. A Declarative Approach to Visualizing Concurrent Computations. IEEE Computer. October 1989.

[Cleveland 93]
W. S. Cleveland. Visualizing Data. AT&T Labs. 1993.

[Gómez 95]
L. Gómez. Sistematización y Uso de las Técnicas de Visualización de Programas Concurrentes. PhD Thesis. DATSI, UPM. 1995.

[Heath/Etheridge 91]
M. Heath, J. Etheridge. Visualizing the Performance of Parallel Programs. IEEE Software. September 1991.

[Keller/Keller 93]
P. R.Keller, M.M. Keller. Visual Clues. IEEE Computer Society Press. 1993.

[Leonardo]
Leonardo (U. di Roma). Home page: www.dis.uniroma1.it/~demetres/Leonardo/

[PV]
PV (IBM). Home page: www.research.ibm.com/pv

[Stasko 90]
J. Stasko. Tango: a Framework and System for Algorithm Animation. IEEE Computer 23(9). 1990.

[Stasko et al. 98]
J. Stasko et al. Software Visualization: Programming as a Multimedia Experience. The MIT Press. 1998.

[Tomás et al. 91]
M. A. de Tomás, L. Gómez, A. Pérez. Vestal: a Tool for Teaching Concurrency in Ada. Proc. of TRI-Ada '91 Conf. ACM Press. 1991.

[Tufte 83]
E. Tufte. The Visual Display of Quantitative Information. Graphics Press. 1983

[Tufte 90]
E. Tufte. Envisioning Information. Graphics Press. 1990

[Tufte 97]
E. Tufte. Visual Explanations. Graphics Press. 1997.

Program Visualization in Theory and Practice

Mordechai Ben-Ari

Visualization and animation of programs has been suggested as a method of improving understanding by novice students. In this paper we survey several research programs by Mayer, Petre and Stasko which show that graphical representations are effective only under certain conditions. We developed the program animation system Jeliot 2000 and carried out a year-long empirical study of its effectiveness. The students showed improved understanding because Jeliot gave them a concrete vocabulary in which to describe the execution of a program. A theoretical explanation of this success is that the concrete vocabulary aids in the development of a viable mental model.

Keywords: program visualization, program animation, mental models, Jeliot.

Introduction

Frank McCourt's recent book Angela's Ashes is a highly-acclaimed memoir of the author's childhood in Limerick, Ireland. With humour and sympathy, he describes the sufferings of his impoverished family from hunger, disease, filth and drunkenness. I eagerly awaited the movie – but I was disappointed. There is no doubt that the movie showed the depressing slum in which McCourt lived and the green fields to which he sometimes escaped, but the movie could not depict the feelings, emotions and thoughts of the characters, and it was precisely these descriptions that made the book a best-seller. Clearly, there are situations in which graphics is not better than text.

The last sentence of the previous paragraph is at odds with conventional wisdom, which holds that graphics is always and necessarily better than text. We have been brainwashed into believing that a graphical user interface (GUI) is obviously better than a non-graphical interface; that graphical representations of programs, such as flowcharts and UML (unified modelling language) diagrams, are better than the textual source program and that colourful animated transparencies are essential for a convincing presentation. Surprisingly, the empirical evidence is not unequivocal and it is difficult to demonstrate significant advantages of graphical representations. How can we resolve this paradox and reconcile the lack of evidence with the intuitive feeling that graphics is better? This paper will review some empirical results obtained throughout the world during experiments with visualization and present theoretical arguments that will help the reader understand exactly what contribution visualization can make. For an introduction to the topic of program visualization, see the collection edited by [Stasko et al. 98].

Empirical studies of graphical representations

Before discussing work on program visualization I would like to review two research programs on the efficacy of graphical representations. Richard E. Mayer of the University of California at Santa Barbara performed a long sequence of experiments on multimedia learning [Mayer 97]. It is axiomatic today that the use of multimedia will improve learning, where multimedia is defined as the use of more than one media. All you have to do is put your lessons on a CD containing text, pictures, sound and video, and your students will learn better. Mayer found that this is not true, and that there are specific conditions that must be fulfilled for multimedia learning to be effective. The presentation of graphics and text must be simultaneous and coordinated. That is, they must appear together and the text must be used to explain the graphics. Mayer's explanation is that multimedia helps students create multiple representations of the problem and these representations must be coordinated. Furthermore, the efficacy of multimedia instruction depends on the characteristics of the learner. Multimedia is more effective for learners with low prior knowledge of the subject and for learners with high spatial ability.

Marian Petre of the Open University in the UK has performed numerous experiments in an attempt to understand the differences between novice and expert programmers and engineers. One set of experiments [Petre 95] focused on the use of graphical representation. She claims that a programmer cannot take in a program like a painting: standing in front of it and receiving a *gestalt*[1] impression of the whole; rather, the pur-

Mordechai Ben-Ari received a Ph.D. in mathematics and computer science from the Tel Aviv University in 1982 with a thesis on temporal and dynamic logics. From 1982 until 1995, he worked in industry as a software manager for real-time computer systems, and held adjunct teaching positions at the Technion and Tel-Aviv University. In 1995 he joined the Department of Science Teaching of the Weizmann Institute of Science as an associate professor. His group develops computer science courses for high school students. Prof. Ben-Ari is the author of six textbooks on concurrency, programming languages and mathematical logic. His latest book: "Mathematical Logic for Computer Science (Second edition)" will be published shortly by Springer-Verlag London. <moti.ben-ari@weizmann.ac.il>
<http://stwww.weizmann.ac.il/g-cs/benari>

to Contents, page 1

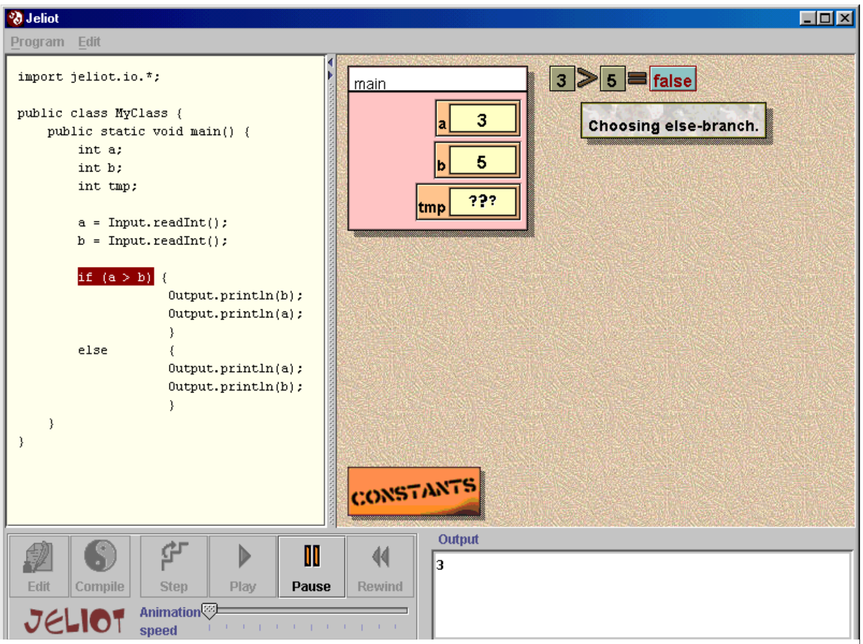


Fig. 1: Screenshot of *Jeliot 2000* showing the animation of an if-statement

pose of a program is to present information clearly and unambiguously, and this requires purposeful examination of the program. Petre's main finding is that the advantage of a graphical representation is primarily in its secondary notation, not in its formal syntax and semantics. Secondary notation includes layout (such as grouping of items), typographic cues (such as indentation), and other graphical enhancements that are part of the conventions used with the representation. These aspects of the notation must be learned and are not obvious to the novice.

In her experiments, Petre found that poor use of secondary notation distinguishes a novice from an expert. The main difficulty that novices have is in determining what is important and what is relevant. Experts have acquired abilities to perceive the relevant information for each task. As in Mayer's experiments, experienced readers used the text to guide their reading of the graphics. Surprisingly, she found that the use of graphical representations was always significantly slower than the use of text for both novices and experts.

Classification of visualizations

A Computer Science (CS) educator seeking to use visualization must be aware of the wide variety of systems that have been proposed, and of the fact that many everyday notations are, in effect, graphical representations. Even something as mundane as conventions on indenting the text of the source program is an attempt to visualize the control structure of the program. Indenting is an example of a static visualization, while another would be a UML diagram that visualizes the structure of an object-oriented design. Animations are dynamic visualizations which attempt to assist the student in understanding the control and data flow of an algorithm or program.

Another dimension of classification is the level of object being visualized. A visualization can be low-level: displaying individual variables, values and statements, or it can be high-level: displaying program modules and data structures.

Clearly, in choosing a visualization system for classroom use, you must take into account your pedagogical objective. A low-level visualization would be appropriate for teaching elementary pointer manipulation, like inserting a new element in a linked list, while a high-level visualization would be appropriate when teaching advanced algorithms like heapsort.

Empirical studies of program visualization

John Stasko of the Georgia Institute of Technology developed the well-known *Tango* and *Polka* algorithm animation systems and has carried out empirical evaluation of the systems. Stasko, Badre, and Lewis [Stasko et al. 93] used algorithm animation to teach a complicated algorithm to graduate students in CS, expecting that the animation would help students understand the dynamic behaviour of the algorithm. The results were disappointing: in a post-test, the group that used animation did not perform better than those that did not. The authors attributed the results to the design of the animations, which was more appropriate for experts than for novices. Novices found it difficult to map the graphics elements of the animation to the algorithm. Kehoe, Stasko, and Taylor [Kehoe et al. 99] evaluated the use of animations in a more realistic setting, and showed that the pedagogical value of algorithm animation is more apparent in open homework sessions than in closed examinations. Furthermore, consistent with Mayer's results, animation is not useful in isolation and students need human explanations to accompany the animations.

Jeliot is a program animation system developed by Jorma Tarhio and Erkki Sutinen of the University of Helsinki.[2] Unlike *Tango* and *Polka*, *Jeliot* is relatively low-level, focusing on program animation rather than algorithm animation. Empirical experiments performed by Lattu, Tarhio, and Meisalo [Lattu et al. 00] showed that the use of *Jeliot* increased motivation, but the four-window user interface (intended to supply the expert user with flexibility in designing the visualization) was too complex for novices (cf. [Petre 95]). Furthermore, the grain of the visualization step turned out to be too large: animation of operations such as loops that are composed of more elementary actions turned out to be confusing. The researchers posed a visualization paradox: if constructing a good visualization requires familiarity with the program, how can a visualization be helpful in understanding unfamiliar programs?

Based on the experience gained with *Jeliot*, we developed a new version, *Jeliot 2000* [Ben-Bassat Levy et al. 00]. The point was not to "fix" *Jeliot*, because the original version was appropriate for its intended purpose, namely, to assist students in the

2. http://www.cs.helsinki.fi/research/aaps

understanding of the source code of algorithms. *Jeliot 2000* was designed specifically for novices. This entailed a design decision: every aspect of the execution of an elementary program was to be animated and labelled (cf. [Mayer 97]) so that nothing would be assumed as "intuitively obvious" (Fig. 1).

We performed a long-term evaluation: Jeliot was used in a full year course in introductory CS taken by tenth-grade high school students, and the students were compared with a control group that received additional instruction without Jeliot. The year-long study enabled the students to become familiar with Jeliot itself, so that we could investigate the contribution of the animation without worrying about the period of time that it takes for a new user to learn to "see" what the system displays (cf. [Petre 95]). The details of the experiment are given in [Ben-Bassat Levy et al. 00]. We can summarize our main conclusion: the animation group used a different and better vocabulary of terms in their explanations and predictions than did the control group.

The students were asked to solve problems that were more advanced than those they had studied in class, such as:

- After learning about assignment statements such as `x:=y+2`, they were asked to predict the outcome of `x:=x+2`, where the same variable appears on both sides of the assignment operator.
- After learning about `if`-statements, the students were asked to predict the outcome of the execution of a nested `if`-statement.

In both cases, some students refused to work on the problems, claiming that the statements were simply incorrect! However, many of the students in the class using *Jeliot 2000* were able to solve the problems by imitating the animation, drawing the trajectories of the values to and from memory, and constructing the labelled evaluation of the `if`-statement. Curiously, some of the stronger students had difficulties answering this question because they believed they could solve the problem without using the animation.

We are currently investigating the teaching of object-oriented programming and Java to tenth-grade students of introductory CS. We are using *BlueJ*,[3] which is a unique development environment intended for teaching. *BlueJ* provides a static visualization of classes and objects, and allows students to work directly with the graphical representation. You can create an object by clicking on the icon for its class and execute a method by clicking on the icon of an object. My first instinct was to begin the course by demonstrating (simplified) GUI elements like frames and labels, in the hope that this would motivate the students. Unfortunately, this did not work as the students conflated the rendering of the GUI elements with the *BlueJ* visualization of the corresponding objects, and conflated the text in the elements with the names of the objects as visualized by *BlueJ*. Again, this shows that visualization can be harmful unless the students are taught what to see.

3. http://www.bluej.org

Theoretical basis of visualization

Is there any reason to suppose that visualization should improve learning? One possible theoretical justification for the use of visualization is that it can help construct mental models of abstract phenomena. A mental model [Gentner/Stevens 83] is a cognitive structure used by a person to represent knowledge about a real-world artifact or phenomenon. The mental model can be run in order to explain behaviour in the real world or to predict behaviour. For example, a mental model of force and motion can be used to predict the amount and direction of force that need to be applied to a basketball so that it will fall into the basket.

Mental models are often associated with constructivism, a theory of learning that holds that learners construct meaning for themselves. (See [Matthews 98] for a collection of articles by both proponents and opponents of constructivism.) According to constructivism, learning is accomplished when the learner is exposed to experiences that require him or her to modify the cognitive structures (mental models) previously existing. If we assume that human cognition is not entirely (or even not primarily) verbal but visual, then visual experiences should be effective in encouraging learning. Furthermore, as many phenomena are not only visual, but dynamic, animation should also be an effective aid to learning.

But just to claim that mental models have visual and dynamic components does not justify the indiscriminate use of visualization and animation. For the use of such a system to be effective, the teacher must (a) have a clear idea of the existing mental model of the student, (b) specify in advance the characteristics of the mental models that the instruction is intended to produce, and (c) explain exactly how the visualization will be instrumental in effecting the transition. This is an immensely difficult undertaking. First, the existing mental models of the individual students are different and it takes quite a lot of effort to elicit even an approximation of a cognitive structure. Second, the goals of learning are usually expressed in terms of "covering the material" of a subject, rather than as characteristics of a desired mental model. Finally, the effects of using a visualization system can only be determined by long-term research that few CS educators have the capability, resources or even the inclination to perform.

Theoretical explanations

In this section I will give two examples of how theory can be relevant in the design and use of visualization. The first is taken from the ubiquitous GUIs that all modern software packages have, and the second is from the use of visualization to present a model of a computer when learning introductory CS.

[Turkle/Papert 90] claim that icons can empower large segments of the population to use computers. However, even with GUIs, computers are still not considered user-friendly for all but the simplest applications. The problem is that an icon is just a representation that will be useful only if the user can construct a mental model of the object being represented. From the constructivist point of view, what is important is the construction of the model and not the sign that denotes it. The use of an icon does not automatically facilitate this construction, and may even cause confusion. For example, suppose that a GUI uses an icon of a bottle of paste to represent the paste operation on a document. The first step that the user must perform is to decipher that the icon represents a bottle of paste and not, say, a bottle of ink or a rocket. Once this is done, the user still has only the word "paste", which in his or her existing mental model is likely to mean "form a permanent chemical bond between one item and another." We should not be surprised if the learner cannot construct a viable [von Glasersfeld 89] mental model and successfully perform the paste operation which means "insert a copy of the material held in an internal buffer into the current working document at the place pointed to by the cursor". In fact, icons seem to have the characteristics of a secondary notation (to use Petre's terminology) that is more appropriate for experts than for novices! This could explain the popularity of tool tips, textual labels that pop up when you place your mouse over an icon. We see that icons are not very relevant to ease of learning as the student still has to learn (construct) the underlying concepts.

In my paper on constructivism in CS education [Ben-Ari 98], I claim that the primary source of difficulty in studying introductory CS is that students do not have a mental model of a computer on which to construct further knowledge. This is consistent with [Smith et al. 93] who claim that you can not learn if you do not have an existing model, even though the existing model is probably non-viable. The theoretical basis for the development of *Jeliot 2000* described above is that a comprehensive animation of the basic operations of a computer will be invaluable in assisting learners of introductory CS by enabling them to build a viable mental model of a computer. I am encouraged by the preliminary positive results of the empirical experiment, namely, that *Jeliot 2000* provided a vocabulary that became part of the learners' mental models and enabled them to run the models in order to explain and predict phenomena. This is consistent with sociolinguistic versions of constructivism which hold that verbalization is the first and most important step in understanding a concept.

Conclusion

Visualization is not a panacea: the mere fact that you visualize something will not necessarily lead to improved learning. Nevertheless, visualization can be extremely important if carefully used:

- The use of visualization just to improve motivation can be justified, especially in CS education, but don't expect miracles if this is all that can be accomplished.
- The operation of a visualization system is not obvious; the user must be taught to see what is displayed by the visualization.
- You must allocate additional time both to teach the visualization and to work with the visualization.
- The visualization must be designed for a specific class of learner and depends on both their experience and their learning styles.
- There must be a theoretical explanation what exactly the visualization is intended to achieve and how it will achieve it.

I believe that as additional empirical research on visualization is performed, we will gradually develop a deep understanding of the potential contributions of visualization in CS education, and of the design principles that must be adhered to in order to achieve these goals.

References

[Ben-Ari 98]
Mordechai Ben-Ari. Constructivism in computer science education. SIGCSE Bulletin. 1998, 30 (1):257–261. Expanded version to be published in The Journal of Computers in Mathematics and Science Teaching.

[Ben-Bassat 00]
Ronit Ben-Bassat Levy, Mordechai Ben-Ari, and Pekka A. Uronen. An extended experiment with Jeliot 2000. Program Visualization Workshop, Porvoo, Finland.
http://stwww.weizmann.ac.il/g-cs/benari/articles/jel.pdf

[Gentner/Stevens 83]
Dedre Gentner and Albert L. Stevens, editors. Mental Models. Lawrence Erlbaum Associates, Hillsdale, NJ, 1983.

[Kehoe et al. 99]
Colleen Kehoe, John T. Stasko, and Ashley Taylor. Rethinking the evaluation of algorithm animations as learning aids: An observational study. Technical Report GIT-GVU-99-10, Georgia Institute of Technology.
ftp://ftp.cc.gatech.edu/pub/gvu/tech-reports/1999/99-10.pdf

[Lattu 00]
Matti Lattu, Jorma Tarhio, and Veijo Meisalo. How a visualization tool can be used: Evaluating a tool in a research and development project. In 12th Workshop of the Psychology of Programming Interest Group, 2000. pages 19–32, Corenza, Italy.
http://www.ppig.org/papers/12th-lattu.pdf

[Matthews 98]
Michael R. Matthews, editor. Constructivism in Science Education. Kluwer Academic Publishers, Dordrecht. 1998

[Mayer 97]
Richard E. Mayer. Multimedia learning: Are we asking the right questions? Educational Psychologist. 1997, 32(1):1–19.

[Petre 95]
Marian Petre. Why looking isn't always seeing: Readership skills and graphical programming. Communications of the ACM. 1995, 38(6):33–44.

[Smith et al. 93]
John P. Smith III, Andrea A. diSessa, and Jeremy Roschelle. Misconceptions reconceived: A constructivist analysis of knowledge in transition. The Journal of the Learning Sciences.1993, 3(2):115–163.

[Stasko et al. 93]
John T. Stasko, Albert Badre, and Clayton Lewis. Do algorithm animations assist learning? An empirical study and analysis. Proceedings of the INTERCHI '93 Conference on Human Factors in Computing Systems. 1993, 61–66.

[Stasko et al. 98]
John T. Stasko, John Domingue, Marc H. Brown, and Blaine A. Price, editors. Software Visualization: Programming as a Multimedia Experience. MIT Press, Cambridge, MA. 1998.

[Turkle/Papert 90]
Sherry Turkle and Seymour Papert. Epistemological pluralism: Styles and cultures within the computer culture. Signs: Journal of Women in Culture and Society, 1990, 16 (1):128–148.

[von Glasersfeld 89]
Ernst von Glasersfeld. Cognition, construction of knowledge and teaching. Synthese, 1989, 80(1):121–140. Also in [Matthews 98].

to Contents, page 1

A User-Centred Approach for Designing Algorithm Visualizations

Sami Khuri

Advances in computing technology and the affordability of software and high-performance graphics hardware enabled rapid growth of visual tools. Today, not only very expensive workstations, but also low-cost PCs are capable of running computationally demanding visualization systems. Algorithm visualizations or the graphic depictions of algorithms in execution are being used in explaining, designing, analysing algorithms, and in debugging, fine-tuning, and documenting programs. Although many tools have been developed over the past twenty years, little attention has been paid to the analysis of users, their needs, tasks, and goals. This paper gives a brief overview of the preliminary design stage of algorithm visualizations, namely the analysis of requirements.

Keywords: Algorithm Visualization, Educational Software, Visual Representation of Information, Design Guidelines, User Centred Approach

1 Introduction

Visualization is defined in the dictionary as "mentally visual images". In the field of computer science the term has a more specific meaning: "The technical speciality of visualization concerns itself with the display of behaviour, and particularly with making complex states of behaviour comprehensible to the human eye" [Gallagher 95]. The term was first made popular in the USA in a 1987 National Science Foundation initiative on scientific visualization that provided first definitions, goals and examples of visualization in scientific Computing [McCormick et al. 87]. Since then, numerous applications ranging from the visualization of mathematical and scientific data to the simulation of virtual environments have been developed. In computer science education, visualization tools can help instructors in a variety of ways, ranging from merely attracting students' attention to increasing students' understanding. Visualizations have been used in teaching, designing and analysing algorithms, producing technical drawings, debugging, fine-tuning, and documenting programs.

Many new algorithm visualizations are being developed each year, yet descriptions of visualization systems rarely specify any particular task they were intended to support [Petre et al. 98]. Creating educational algorithm visualizations requires substantial time and effort. Mapping an algorithm to an animated representation is a non-trivial problem; it requires careful thought and knowledge of a particular algorithm animation programming framework. Today's IDEs[1] provide many gadgets for designers to experiment with, and the emergence of the Java programming language allows them to make their systems available over the Internet. It is so easy to pull down menus and select different fonts, assign vivid colours and embed an applet in the Web page. Developers of educational visualization packages claim that their systems can be used by all kinds of users, novice and experienced, and in all kinds of tasks: demos, homework, laboratories, and self-study. But, no algorithm visualization will ever be universally superior across all kinds of users and tasks. In what follows, we describe a "user-centred" framework for designing algorithm visualizations, in which visualizations are developed for real needs, real users and real tasks, and not just because the technology to implement them is there.

1. IDE: Integrated Development Environment

2 Designing Algorithm Visualizations

Like many other design disciplines, a successful algorithm visualization design should consider the many facets of design issues such as effective representation and presentation, objectives, environmental considerations, design and layout, colour, graphics, and user interface. The process of producing algo-

Sami Khuri holds a M.Sc. degree in Mathematics, a M.Sc. in Computer Science and a PhD. in Computer Science, all three degrees from Syracuse University, USA. He is a Professor of Computer Science at San Jose State University, USA. He was the recipient of several research awards. He was a Dana scholar in the Biomedical Engineering Department at the Johns Hopkins School of Medicine, a Fulbright scholar at the American University in Bulgaria and a DAAD scholar at the Technical University of Munich in Germany. His research interests centre around the design and analysis of algorithms, genetic algorithms, visualization and algorithm animation. He published and presented his work at international Computer Science conferences and gave tutorials and lectures on Data Compression, Genetic Algorithms and Algorithm Visualizations. Sami Khuri has been involved in the design and implementation of interactive algorithm-specific visualizations used in Operating Systems, Compiler Design, Data Compression and Neural Network courses.
<khuri@cs.sjsu.edu>,
<http://www.mathcs.sjsu.edu/faculty/khuri/index.html>

Part of this work was done during author's sabbatical leave at the Departamento de Lengajes y Ciencias de la Computación, Universidad de Málaga.

rithm visualizations is time-consuming and resource consuming and includes at least the following five steps: analysis of requirements, system design, implementation, testing, and maintenance. Obviously, these stages will often overlap and only a few systems pass smoothly through these steps. Testing, for example, may reveal ambiguities in the design specification, parts of which will have to be rewritten. Analysis of the requirements is always the most important stage, for it is the one that justifies the other four. Unfortunately, it is often the one that is skipped. For education software visualizations, this stage involves trying to provide questions, such as: Who will use the system? How would it fit into existing curricula? What would the system add to what could be done by other means? Is the system technically feasible? In what follows, we bring to the readers' attention the important parts of the analysis of requirements.

2.1 Users' Unalysis

Understanding who the users are determines the system's content, organization, breadth, depth, and information presentation. Users of algorithm visualizations can be categorized into four roles: student, educator, researcher or developer (it is possible that one individual will assume more than one of these roles). Another possible approach is to characterize users into novice and experienced users. It is difficult to develop a system suitable for both novice and experts. The needs of programmers of different levels of expertise may not be satisfied by a single system. Users at the beginning levels do not learn well by trial and error (i.e. do not profit from "floundering" and trying to find their own way to correct paths after following incorrect ones for some time). They tend to have problems in mapping the real world model onto a program and they are easily distracted by motion and by extraneous detail. They easily form misconceptions and need as many of the following design features as possible:

- Graphical and consistent means of visualization control (e.g. menus, buttons, etc).
- Similar instructional structure of all visualizations, such as a number of worked examples showing how the algorithm can be used (build-in default visualizations), as well as a graphic support tool for working with new problems (e.g. the facility to input new data sets, change parameters of the algorithm, etc).
- Comprehensive help files with clear directions on how to use the tool, descriptions of the algorithm's steps, as well as the organization of the interfaces.
- Short "quizzes" or similar opportunities for students to evaluate whether they have understood the material, and to exercise the use of the visualization tools.

Expert users on the other hand, will want to "play" with the code, see how it works, and then modify, expand or otherwise just experiment with it. For example, they might want to see how they could integrate these modules with other tools. Thus, animations designed for experts might require the ability to move between algorithm-level and program-level displays depending on their needs.

2.2 Needs Analysis

After determining the group of visualization's users, it is very important to address their needs. To do so, it is necessary to realize the fact that different students have different learning styles and strategies. Several factors such as locus of control, motivation, and learners' expectation, all play a significant role in learning. Some students prefer the "hands on" approach, where they are actively involved in the learning process. These students tend to explore more and are generally more independent learners. However, some students prefer to be led through the lesson, allowing the instructor, or the computer to control the flow of the material. For the latter type of users, algorithm visualizations should graphically demonstrate effects of the algorithm on the data structures. It is also useful or even necessary, to support the animation with explanatory cues in the form of short, on-screen, textual notes, and to provide control of the speed at which the algorithm is animated.

Support for users who prefer active interaction requires a tool-set similar to that found in program development environments. In particular, the availability of "debugging" style software to allow single step execution, breakpoint setting and the monitoring of key variables is essential. The ability to simultaneously view the algorithm execution path and the data structure state is a crucial aspect of interactive observation.

Before implementing an algorithm visualization system, the designer should also ask the following question: Do the users need this information presented in this way? For example, an instructor can illustrate the binary search algorithm by using a phone book. In general, effort is wasted on visualizing simple concepts or algorithms. For example, if the amount of data is small, or the data structure is very simple, or the relationship of objects is important but movement is not needed, then it is better to use a static picture. But for a large amount of data, and for complex data structures or when movement is needed to show how the relationships between objects change over time, animation is the right choice for the presentation of the algorithm.

2.3 Task Analysis

When designing an algorithm visualization system, it is important to note the system's intended goals and select the content accordingly. The user might use the system to create new animations, interact with existing visualizations to understand the behaviour of an algorithm, or visually debug programs. Each situation demands a different kind of visualization system, and it is difficult to build the system that satisfies all of them.

Traditionally, computer science instructors constructed visualizations that were later used either as visual aids in lectures, such as *BALSA* [Brown 88] or in closed laboratories, such as *GAIGS* [Naps 90]. More recently, computer science educators have advocated using algorithm visualization software, such as *XTango*, as the basis for visualization assignments, in which students construct their own visualizations of the algorithms under study [Stasko 97].

If a system is designed for classroom teaching, it might intentionally show algorithm-level displays only and avoid program-level displays in order to keep the student's minds off implementation details. Systems designed for user's exploration might require the ability to change settings, input different data sets, change speed, move one step back, or move between algorithm-level and program-level displays.

If novice users are expected to use the system to design their own visualizations, and if making animations is difficult and tedious, then their effort and time is wasted in low-level graphics programming. In one study, students spent over 33 hours on average constructing single visualizations with *JSamba* [Hundhausen 99]. The system designed for this purpose should have a powerful editor allowing students to map graphical objects to data structures automatically.

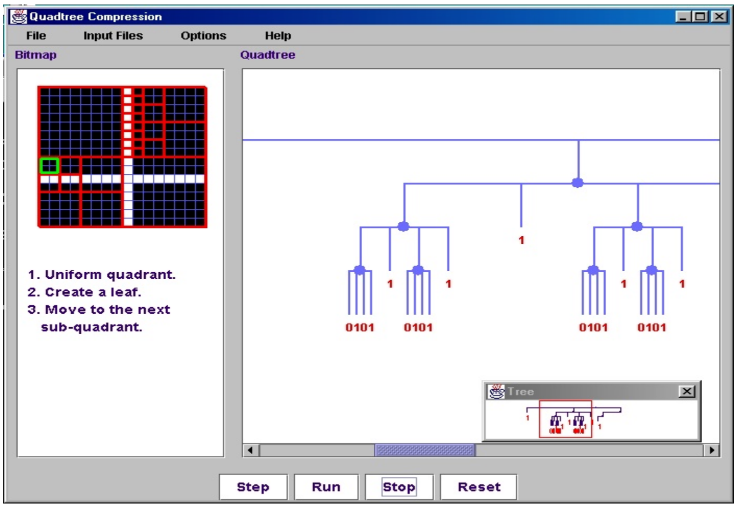


Fig. 1: Snapshot of the Quadtree package [Khuri/Hsu 00].

2.4 Information Analysis

Information should be analysed to determine how effectively it could be visualized. Viewers should be able to forget about the technique of presentation and concentrate instead on what is being taught. Without an appropriate set of visual conventions, such as one colour to denote some items and another for other items, one may spend more energy trying to figure out what the picture means than in trying to follow the algorithm.

The graphical representation of information is heavily dependent on the concept we would like to visualize. For example, an important learning objective for students in understanding the bubble sort algorithm is to first comprehend the central metaphor of large values "bubbling up". Another learning objective is to understand the core operation of swapping two data values that sorting algorithms employ. It is often useful to provide multiple views of the same system in order to understand a variety of characteristics of the data. Multiple views might include a graphical view of changing program data with a corresponding view of the executing source code. For example, the package shown in Figure 1 animates the quadtree compression algorithm for bitmap images. The quadtree method scans the bitmap, area by area, looking for areas filled with identical pixels. The package takes a bitmap as input, and constructs a quadtree, where a node is either a leaf or has exactly four children. The construction of the tree is done using a very interesting recursive algorithm. The area on the left-hand side of the application (see Figure 1) is used for displaying the bitmap image and steps of the algorithm. The red lines show how the quadtree algorithm partitions the image. A green rectangle indicates the quadrant being evaluated by the algorithm. The window on the right entitled "Quadtree" shows the building of the tree. A useful extension to the multiple windows approach is that of semantic zooming. The tool named "Tree" displays the whole quadtree but on a smaller scale than the one depicted in the "Quadtree" window. The current view of the "Quadtree" window is always the part of the tree inside the red rectangle in the display tool. Users can view other parts of the quadtree in the "Quadtree" window by simply moving the red rectangle to the desired location in the display tool.

In general, there are two approaches to visualizing different, but related algorithms. With the unified-view visualization, one graphical representation is used for different algorithms of the same class, and designers try to keep the individual adaptations as small as possible. The advantages of this approach is time saving since once the animation view has been established for the first algorithm, the views can then be reused. Because of the common base, the behaviour of related algorithms could be compared more easily. Another possible way of animating related algorithms is to use a unique view. In this approach, a unique representation is developed for each algorithm. It takes more time and effort to visualize algorithms this way because the best representation must be found for each algorithm, and the comparison of uniquely-visualized algorithms requires additional work, but the advantages of this method are meaningful graphical metaphors and increased memorability of the visualization. Care should be taken in selecting graphical metaphors. A given picture can mean several different things to different viewers, and the meaning will change depending on the relation to other simultaneously displayed images.

Developers have several choices of graphic representation [Cox/Roman 92]. In direct representation, data structures of an algorithm are directly mapped to a picture (e.g. representing an array as a collection of objects, where the index of each array element is mapped to the object's X-coordinate, while the value of the element is mapped to the object's Y-coordinate). In structural representations, some details or information are hidden and the remaining information is directly represented. For example, upon visualizing a computation running on a network of processes, we might want to reduce the complex states of the individual processes and simply show them as being in one of two states: "active" or "inactive". In this representation, we conceal the other attributes of the processes. Some representations can be of synthesized nature, i.e., the information of interest can be derived from the program data, but is not directly represented in the program. Some examples include counting the number of items already sorted or compression ratios. A slightly different type of information representation is that of explanatory nature. In this type of representation, visual events have no counterparts in the underlying algorithm. They are added to enhance the presentation. Their goal is to communicate the implications of a particular computational event or to focus the viewer's attention.

Some of the many ways of representing information is by using shape, size, colour, texture, and arrangement of objects, sound, and 3D. Colour can be used to call attention to specific data, identify elements or structures, depict logical structure, increase the number of dimensions in coding the data, and highlight relationships. Although, colour can be a very powerful way of representing the information, some caution must be taken. With respect to colour, it seems best to be conservative. Only four distinct colours should be appropriate for novice viewers. This allows extra room in short-term memory (about 20 seconds), which can store up to five words or shapes, six letters, seven colours and eight digits. The same colour should be used for grouping related elements. It is important to be complete and consistent. For example, command and control colours in menus should not be used for information coding within a work area unless a specific connection is intended. Similar background colours of related areas can orient the viewer to understand the conceptual linking of the two areas.

Sound is another interesting area of research in information visualization. It is a useful complement to visual output because it can increase the amount of information communicated to the user or reduce the amount of information the user has to receive through the visual channel. Although not suitable for conveying exact values, auralization can indicate trends and increase the number of dimensions capable of being presented simultaneously.

2.5 Scope Analysis

The scope can range from single-purpose visualizations that illustrate one algorithm or a group of related algorithms in detail, to specialized systems that concentrate on algorithms in certain fields of computer science, such as graph algorithms and finally, to general purpose systems. The latter can (ideally) animate any algorithm. The greater the number of algorithms that can be animated, the more desirable the result. See [Khuri 00] for a collection of links to different algorithm visualization systems. Before embarking on developing a general-purpose system, one should consider that increased flexibility results in increased complexity. In general, systems that restrict themselves to animating only algorithms in one field, such as geometrical algorithms, might not be able to easily represent algorithms in other fields. On the other hand, specialized packages can be polished to make pleasing, informative visualizations of frequently used objects. Some of the recommendations for the designers are:

- Design small first. Don't attempt to provide everything possible in the beginning. Provide what you can that is beneficial to the user, visually attractive, and is of high quality.
- Plan a phased growth. The visualization might grow and change over time. Make sure to plan for growth by using object-oriented design and carefully documenting the programs. Some of the algorithm-specific visualizations might not allow adding new features. Plan to add new features over time and make upgrades publicly available.

2.6 Resource Analysis

Remember that design takes longer than expected. Designing visualizations is not simple, especially when the designer is concerned with more than just the visualization appearance. Visualizations places great demands on computer resources, such as CPU speed, monitor size and resolution, RAM and disk memory sizes, networking, audio and video input and output, colour display panel and projection, and other input and output equipment needed. If visualization will be available over the Internet, they will need large amounts of storage space, and more importantly, a high-speed connection to the net.

Another purpose of the resource analysis is to select the specification method for designing new algorithm visualizations. Some of the possibilities are annotation, declaration, manipulation, and predefinition. In an annotation method, important steps of an algorithm are annotated with interesting events. The interested events call graphical operations which execute animations (move rectangles, change colours, etc.). When these program points are reached during execution, events are created and then forwarded to different views of the algorithm animation system. These views represent the interesting events by appropriate animations, as in *BALSA* [Brown 88]. In *PAVANE* [Cox/Roman 92], for example, the user can specify a mapping between the program's state and the final image arbitrarily or by declaration. The changes in the state will be immediately reflected in the image. This approach provides greater abstract capabilities, but this increased power also requires more processing to map the program to the final image. One of the interesting, but rarely found systems, is the one that allows the creation of new visualizations through manipulation (or animation by demonstration). The user can specify visualizations through the use of examples. The system attempts to capture the gestures used by the animator as she directly manipulates an image and ties these gestures to specific program events. This approach suffers from the difficulty of specifying the exact relationship between the gesture and the program event. The predefinition method is often used for application-specific visualizations, such as Quadtree in Figure 1, and employs a fixed or highly constrained mapping. There is little or no control over what is visualized and in the way the information is presented.

3 Concluding Remarks

In this paper, the preliminary design phase: the analysis of requirements is discussed. Many algorithm visualizations are being designed without paying too much attention to the needs of users, their tasks, and their characteristics. Creating visualizations requires substantial time and effort. Mapping an algorithm to an animated representation is a non-trivial problem; it requires careful thought and knowledge of a particular algorithm animation programming framework. There are many underutilized techniques for creating effective algorithm visualization, such as sound, 3D, and artificial intelligence, which will continue to spark the creativity of the developers in the years to come. But the main problem still remains the same, there is no single algorithm visualization, specialized or general-purpose, that can satisfy all kinds of users, all kinds of tasks and can be used in all kinds of environments. Without the careful analysis of the users' needs, tasks, scope, information, and resources, the effort put in developing a visualization package will probably be wasted. Readers interested in obtaining more information about designing effective algorithm visualizations are referred to [Khuri 00] and [Stasko 98].

References

[Brown 88]
M. Brown: "Algorithm Animation", MIT Press, Cambridge, MA. 1988.

[Cox/Roman 92]
K. Cox and G. Roman: "Abstraction in Algorithm Animation", Technical Report WUCS-92-14, School of Engineering and Applied Science, Washington University in St. Louis, 1992.

[Gallagher 95]
R. Gallagher: "Computer Visualization: Graphics Techniques for Scientific and Engineering Analysis", CRC Press, 1995.

[Hundhausen 99]
C. D. Hundhausen: "Toward Effective Algorithm Visualization Artifacts: Designing for Participation and Communication in an Undergraduate Algorithms Course", Ph.D. Dissertation, University of Oregon, June 1999.

[Khuri/Hsu 00]
S. Khuri and H. Hsu: "Interactive Packages for Learning Image Compression Algorithms", Proceedings of the 5th ITiCSE, 2000, pp. 73–76.

[Khuri 00]
S. Khuri: "Designing Effective Algorithm Visualizations", Invited Lecture, Program Visualization Workshop, Porvoo, 2000, available at http://www.mathcs.sjsu.edu/faculty/khuri/invited.html

[McCormick et al. 87]
B. H. McCormick, T. A. DeFanti and M. Brown: "Visualization in Scientific Computing", Report of the NSF Advisory Panel on Graphics, Image Processing and Workstations, 1987.

[Naps 90]
T. L. Naps: "Algorithm Visualization in Computer Science Laboratories", Proceedings of the ACM SIGCSE, 1990.

[Petre et al. 98]
M. Petre, A. Blackwell and T. Green: "Cognitive Questions in Software Visualization", in Software Visualization, MIT Press, 1998, pp. 453–480.

[Stasko 97]
J. T. Stasko: "Using Student-Built Algorithm Animations as Learning Aids", Proceedings of the ACM SIGCSE, 1997.

[Stasko 98]
J. T. Stasko: editor, Software Visualization, MIT Press, 1998.

Incorporating Algorithm Visualization into Educational Theory: A Challenge for the Future

Thomas L. Naps

The primary goal of this paper is to examine AV and the role it can play in the context of learning models recognized by educational theorists.

Algorithm visualizations (AV) use computer graphics to depict the actions of an algorithm. Many AV tools have been built and are freely available. The reasons for their development are threefold. For researchers in the field of AV, there is the challenge of developing new visualization techniques. For practitioners, AV can help in the process of designing and debugging algorithms. For computer science students, AV holds promise to help them understand algorithms more easily and in greater depth. In 1981, Ronald Baecker's Sorting Out Sorting animation [Baecker 98] started what many hoped would be a revolution in computer science education. Marc Brown's award-winning Ph.D. dissertation on algorithm animation [Brown 87] fueled this fire. Increasingly many algorithm visualization tools have been developed and presented at recent SIGCSE and ITiCSE conferences (eight papers in '98 SIGCSE proceedings, nine in '98 ITiCSE proceedings, ten in '99 SIGCSE proceedings, four in '99 ITiCSE proceedings, seven in SIGCSE 2000 proceedings, and nine in 2000 ITiCSE proceedings). Yet AV remains a pedagogical enigma. In [Baecker 98] Baecker recently lamented "how difficult it still is to describe and control algorithm animations – and how little the work (that researchers have done in algorithm visualization) has been adopted by the mainstream of computer science education and practice."

Too often in the past, according to Petre, Blackwell, and Green [Petre et al. 98] "descriptions of visualization systems rarely specified any particular task they were intended to support." Many of "the current algorithm visualizations concentrate on graphics rather than on pedagogy" [Stern et al. 99]. These tools lack features to encourage students' interaction with the AV system.

The primary goal of this paper is to examine AV and the role it can play in the context of learning models recognized by educational theorists. In [Howard et al. 96], Howard, Carver, and Lane describe their experiences in applying several such strategies in teaching CS2. Among these strategies are:

- Bloom's Taxonomy of Learning
- Schwartz's Minute Paper
- Felder's Learning Model
- Kolb's Learning Cycle

Each of these models will be described in general, and we will then examine the potential of using AV within each model.

Thomas L. Naps received the PhD in mathematical logic from the University of Notre Dame in 1975. He has taught a broad range of mathematics and computer science courses, first in the University of Wisconsin Centre System and then at Lawrence University. He has authored twelve computer science textbooks, focusing on the area of data structures and algorithms. A unique feature of many of these texts is drawings that accompany program listings and depict the actions taken by a particular code segment. His use of graphical documentation in texts spurred an interest in using algorithm visualization (AV) as a pedagogical technique in the classroom. Consequently, since 1987 he has pursued AV both from the perspective of an instructor who wants to design visualizations of particular algorithms for his students and as the developer of two generalized AV software systems. He has authored or co-authored thirteen published papers in the area of AV and conducted workshops on AV under the NSF's Undergraduate Faculty Enhancement Program.
<thomas.naps@lawrence.edu>

1 Description of the Models

1.1 Bloom's Taxonomy [Bell et al. 95], [Bloom 56]:

This taxonomy structures a student's depth of understanding along a linear progression of increasingly sophisticated levels:

Level 1: Bloom terms this the *knowledge level*. It is characterized by mere factual recall with no real understanding of the deeper meaning behind these facts.

Level 2: The *comprehension level*. At this level, the student is able to discern the meaning behind the facts.

Level 3: The *application level*. Now the student can apply the learned material in specifically described new situations.

Level 4: The *analysis level*. The student can identify of components of a complex problem and break it down into smaller parts.

Level 5: The *synthesis level*. The student is able to generalize and draw new conclusions from the facts learned at prior levels.

Level 6: The *evaluation level*. The student is able to compare and discriminate between different ideas and methods. By assessing the value of these ideas and methods, the students is able to make choices based on reasoned argument.

1.2 The one-minute paper:

According to Howard, Carver, and Lane [Howard et al. 96], this idea originated with Charles Schwartz, a professor of physics at the University of California. The essence of the technique is to ask students one or two quickly answerable questions at strategic times during a period of instruction. Students submit the answers to these questions on index cards. Their responses are then examined by the instructor not to necessarily grade the student, but rather to gather feedback on the general progress of the class in mastering a certain concept. Some instructional guidelines for using the one-minute paper appear in [Cross/Angelo 88]. In summary they are:

- Use one or two questions about the content of the material to which you would like your students to respond.
- Ask yourself the extent to which you are willing to act on the students' responses.
- Unless there is a very good reason to know who wrote what, students should respond anonymously.
- Collect and tabulate the data, making note of any useful comments.
- Act on the feedback. Asking students to respond to questions about your course is likely to raise expectations that you are willing and planning to make changes.

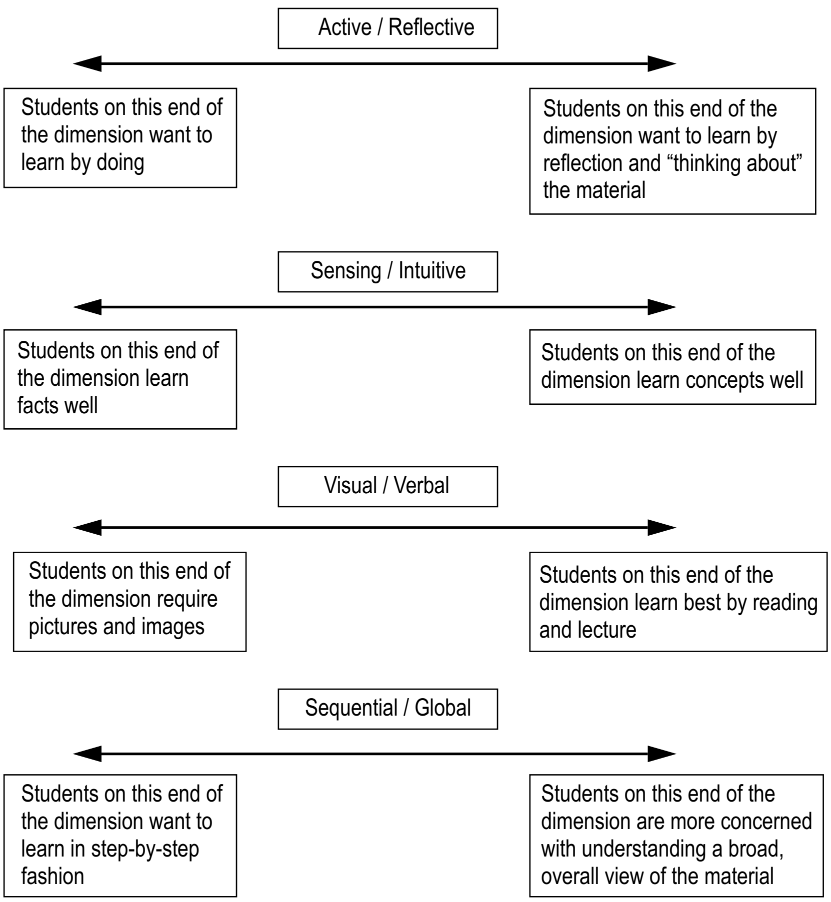


Fig. 1: Felder's four learning dimensions

1.3 Felder's Learning Model and Kolb's Learning Cycle:

Whereas Bloom's taxonomy focuses on the depth of a student's understanding, both Felder's model [Felder 93] and Kolb's cycle [Kolb 84], [McCarthy 87] focus on the variations in the ways in which a student reaches such understanding. Felder portrays a student's learning style as existing along a scale of four different dimensions – active/reflective, sensing/intuitive, visual/verbal, and sequential/global (see Figure 1). According to Felder, there is no "best learning style." Rather it is a question of what style a student prefers in each of the dimensions. To address the needs of all of our students, a teacher will ideally teach at different ends of each dimension at varying times during a course.

The Kolb learning cycle addresses four different learning styles. The first two are (1) reflective observation, in which ideas are examined from several angles but no actions are taken, and (2) active experimentation, in which hands-on experience is involved. These correspond to the extremes of the active/reflective dimension in the Felder model. The last two are (1) abstract conceptualization, in which logical analysis, abstract thinking and systematic planning are preferred; and (2) concrete experience, in which nonsystematic, specific experiences and personal involvement are preferred. These two styles fit into the sensing/intuitive dimension of the Felder model. McCarthy [McCarthy 87] maintains that teachers must consider all four styles in Kolb's model to insure complete learning. Howard Carver, and Lane [Howard et al. 96] refer to this consideration of varied learning styles in our teaching as "teaching around the circle."

2 The Potential Role of AV in these Models

Let's now examine how AV can potentially contribute to a course in which these four pedagogical strategies are incorporated in the teacher's planning. Much of our discussion will centre on our use of JHAVÉ, a Web-based AV environment that we have developed at Lawrence University. [Naps et al. 00] However, other AV systems, such as John Stasko's Tango/Samba [Stasko 90] and Guido Roessling's Animal [Roessling et al. 00], might equally well serve as the AV vehicle for the type of use we describe.

2.1 AV in the context of Bloom's Taxonomy:

2.1.1 Knowledge and comprehension levels:

The levels in Bloom's taxonomy are reflected in the progression a student goes through in understanding an algorithm. The levels Bloom identifies as *knowledge* and *comprehension* correspond to a student's learning the steps of the algorithm as a "recipe" carried out by the computer to solve a particular problem. We initially require that the student understand the steps involved in the algorithm and be able to trace those steps with a typical data set of small size. For example, we might require that the student understand how array values are interchanged in the execution of a particular sorting algorithm. At Bloom's initial *knowledge* level, this understanding may only involve "memorizing" those steps and being able to trace them without understanding why they solve the problem. For more sophisticated algorithms, such as the 0/1 knapsack algorithm, our experience shows that it is entirely possible for a student to learn the technique for filling in the matrix of values without having any intuition as to why those are the correct values. At the *comprehension* level, we hope that the student also acquires a sense of why the algorithm works to solve the problem.

The role AV can play in helping the students achieve these two initial levels of understanding is to provide an "infinite" supply of data sets on which they can watch the algorithm execute. For example, in our JHAVÉ AV system, when the student requests a visualization of a particular algorithm, the JHAVÉ client contacts the JHAVÉ server. The server will execute an instance of the algorithm that produces a script, which is in turn transmitted back to the client to be rendered for the student. Hence the AV system must be much more than a mere renderer; it must also know how to generate instances of problems that will provide instructive content to students. In this regard the Java *problet* interface suggested by Kumar in [Kumar 00] offers tremendous potential. Although it has not yet been incorporated in AV systems, doing so is certainly a worthy area of exploration.

If AV systems are to be truly effective for students working at these lower levels of the Bloom hierarchy, they must present their users with an interface that allows "rewinding" the algorithm's execution. That is, when a student becomes lost or confused in watching a visualization, she must be able to backtrack to the point where she became lost. This was borne out in a study done by Stasko and Lawrence [Stasko/Lawrence 98] in which students were quizzed on a pairing heap algorithm after studying it with the help of the XTango AV system. According to Stasko and Lawrence, "the most often cited negative comment (on the part of the participants) was the inability to rewind the animation. The participants said that, after an operation occurred, they often wanted to look at the heap as it appeared before the operation."

2.1.2 Application level:

As one progresses through increasingly sophisticated levels in the Bloom taxonomy, it becomes incumbent upon the AV system to offer a more sophisticated interface to the student. At Bloom's application level, the student is expected to apply learned material in new situations. With respect to understanding of algorithms, this could mean the ability to detect how the algorithm would perform on certain pathological data sets illustrating best and worst case behaviour. For example, how does the algorithm react to data sets that are not "typical", such as an array that is already in order for a version of quick sort that selects its pivot from the low index? Here, rather than having the problem instance generated by the AV system, we would want the student to ponder and design her own data sets for the algorithm. In the JHAVÉ architecture, we provide *input generators* for this purpose. Each input generator consists of text boxes or pull down menus, depending on the input data the visualization requires. We must be careful that facilities to input data do not overwhelm the student. For example, designing strategic input data for a graph algorithm can be a very time-consuming task. Consequently data input facilities must be carefully constructed to make it necessary for the student to only decide those issues that affect the type of understanding we are trying to achieve. In sorting algorithms, this might mean merely allowing the student to choose between different categories of data sets (for example, randomized, in-order, in-reverse-order, almost-in-order) as opposed to forcing the student to input each value in the array to be sorted.

2.1.3 Analysis and Synthesis Levels:

An example of algorithm understanding at this level of the Bloom hierarchy would be understanding the relationship of the algorithm to its implementation in code. We would expect that the student is able to actually write a program that correctly implements the algorithm. The ideal for this usage of AV is to have the visualization appear as part of a programming environment. The student can write code to implement a complex algorithm and immediately see, in a graphical view, the effects of executing that code. One system designed with this philosophy in mind is Ross's DYNALAB [Birch et al. 95]. This system is more accurately termed a *program animator* then an AV system. The difference here is that a program animator gives a very detailed view of what is happening inside the computer itself when a program executes. Its level of abstraction is not far removed from a sophisticated debugger. AV systems, on the other hand, present a more conceptual view of the algorithm's execution. The linkage of program animators to a more machine-based view of a program's execution can make it difficult for them to provide insightful views of more complex data structures such as trees and graphs. Much work remains to be done in developing such environments. Certainly one of the drawbacks to developing such environments is that they become tied to particular programming languages. Hence, as computer science educators keep switching languages, program animators can become quickly obsolete. For example, Ross had to port DYNALAB from Pascal to C++ and now is faced with a similar dilemma because of the increasing popularity of Java.

2.1.4 Evaluation level:

Algorithmic understanding at this level of the Bloom hierarchy involves the ability to formally derive an algorithm's efficiency and prove its correctness. Here a visualization alone may not be of much help unless it is accompanied by accompanying hypertext material that explains to the student the significance of events that occur in the algorithm she is watching.

This accompanying material can be tied to the animation in three ways that demonstrate successively higher levels of interactivity. We term these three levels *static*, *algorithm-sensitive*, and *dynamic*. In static linkage, the same non-changing hypertext material is always viewable, regardless of the stage of execution of the algorithm. In linkage that is algorithm-sensi-

to Contents, page 1

tive, the hypertext material being viewed changes with respect to the stage of execution of the algorithm. For example, in watching a splay tree animation, the hypertext material available during an LR-rotation would be different than the material available during an RR-rotation. In linkage that is dynamic, the hypertext material is aware not only of the stage of execution of the algorithm but also of the particular data values that are being manipulated. For instance, whereas algorithm-sensitive material might only inform the viewer "the red node will be moved up a level in the tree in this LR-rotation," it will not be aware enough to say "the red node containing 4 will be moved up a level in the tree in this LR-rotation." This is because the text has been written before the execution of the algorithm that it tries to explain. It would be even more valuable if the textual material could be dynamically produced as the algorithm executes. It would then have much more awareness of the data being manipulated and could offer a much more specific explanation to the student. Newer Web delivery techniques, such as Java servlets, offer much potential in the area of dynamic production of hypertext to accompany visualizations.

2.2 AV in the context of the minute quiz

The minute quiz serves two purposes. First, it makes students stop and think about material to which they have just been exposed. Second, the collective results of a class on a minute quiz provide valuable feedback to an instructor about how well material has been understood. With the JHAVÉ AV architecture, we have been able, in effect, to insert numerous minute quizzes into a visualization watched by a student. These quizzes come in the form of quickly answerable questions that pop up at an interesting event during the algorithm's execution. (Figure 2.) Such questions make the student predict what she will see in the next step of the visualized algorithm. Without such questions, once a student becomes confused, continuing to watch a visualization is akin to watching a movie in which one has lost interest. The insertion of questions changes this dramatically. When a confused student answers a question incorrectly, continuing with the visualization not only provides the student with the correct answer but serves to reset the student's perception of the algorithm back on the track intended by the instructor.

In addition to allowing us to focus a student's attention at key points in a visualization, when a student feels that she has mastered the algorithm, she can switch into JHAVÉ's "quiz-for-real" mode. Her responses to questions are then recorded by the JHAVÉ server in a JDBC-accessible database. Instructors can consult this database before a class starts to gain insight on the progress of the students in understanding an algorithm. Our typical usage of JHAVÉ in this regard is to assign students reading about an algorithm for which they should have achieved understanding at both the knowledge and comprehension levels (in the Bloom taxonomy sense) before the next class. There is also the expectation that students will engage in a JHAVÉ-generated minute quiz on the algorithm while watching a visualization of it. We check how students have done on these quizzes before the next class starts, and that

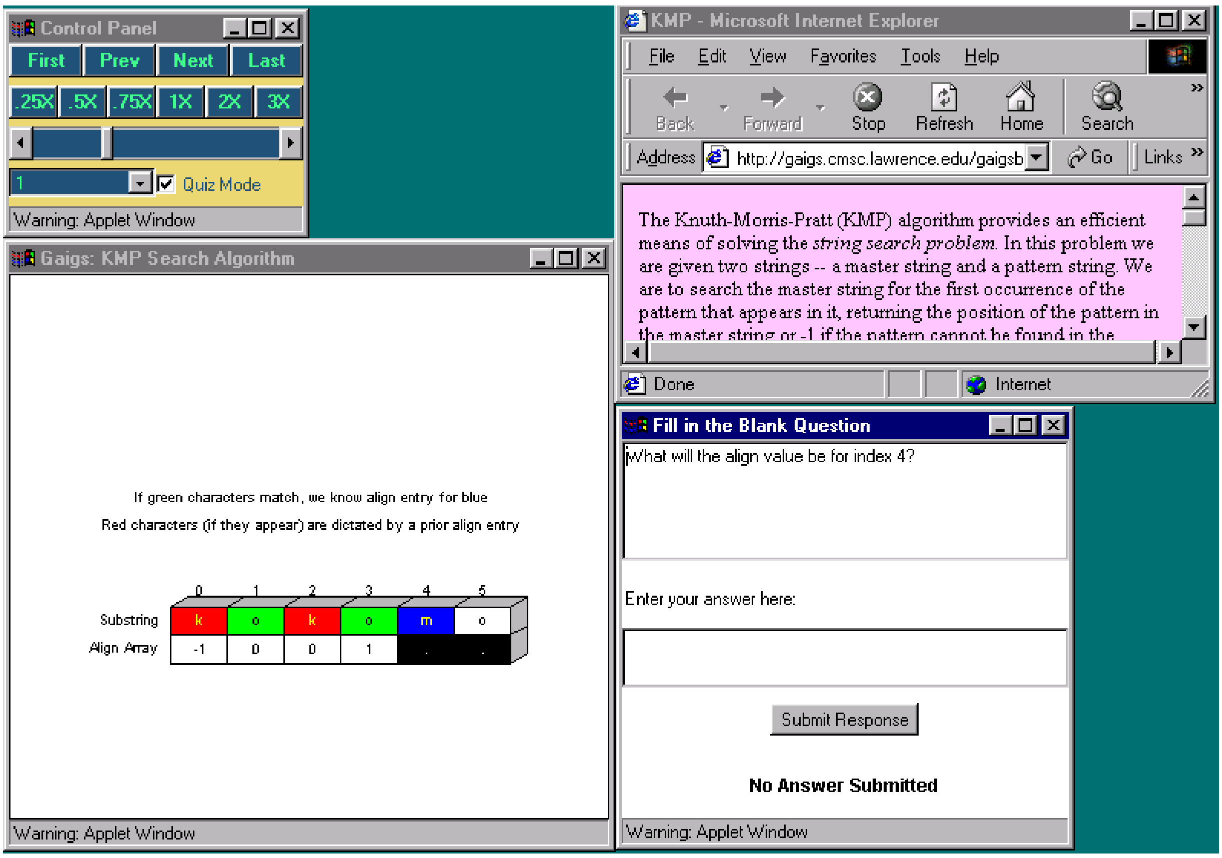


Fig. 2: JHAVÉ's portrayal of the Knuth-Morris-Pratt algorithm with context-sensitive documentation and a fill-in-the-blank question.

gives us an indication of the level at which we should teach the class. In most cases, we've found that students arrive in class having already achieved Bloom's Level 1 and 2 learning objectives for the algorithm. This allows us to not waste valuable class time getting students to these levels. Instead, we can focus our class discussion at the higher levels in the Bloom hierarchy. We thus are able to cover more material than we have in the past without any sacrifice in students' understanding of the material.

2.3 General Conclusions – AV in the context of Felder's Learning Model and Kolb's Learning Cycle

The conclusions of Kolb and Felder tell us that algorithm visualization will be used to its greatest advantage when it is compliments other teaching methods – it cannot be the sole vehicle we use to deliver content about algorithms to our students. However, Kolb and Felder also tell us that we must teach "around the circle." That is, our delivery systems for presenting algorithms to our students must be many and varied to mesh with the dichotomies in our students' learning styles. Students whose learning styles tend to be on the left side of the four dimensions in Felder's model may well learn an algorithm better from a visualization than they will from hearing an instructor talk about it.

The reflective observation and concrete experience components of Kolb's learning cycle are similarly areas in which students may be very well served by algorithm visualizations. If we, as instructors, weave AV into the fabric of our courses, Kolb and Felder leave little doubt that our students will learn better. AV can become an important component of our teaching around the circle if we structure learning experiences around it. We would do well to heed the advice of Bazik, Tamassia, Reiss, and van Dam at Brown University, the site of significant pioneering work in AV. In [Bazik et al. 98], they write, "Much of the success of the BALSA system at Brown is due to the tight integration of its development with the development of a textbook and curriculum for a particular course. BALSA was more than a resource for that course – the course was rendered in software in the BALSA system."

References

[Ann Arbor]
M. I. Ann Arbor: Board of Regents for the University of Michigan, p. 148–150.

[Baecker 98]
R. Baecker: "Sorting Out Sorting: A Case Study of Software Visualization for Teaching Computer Science," in Software Visualization: Programming as a Multimedia Experience, The MIT Press 1998, pp. 369–382.

[Bazik et al. 98]
J. Bazik, R. Tamassia, S. P. Reiss, A. van Dam: "Software Visualization in Teaching at Brown University," in Software Visualization: Programming as a Multimedia Experience, The MIT Press 1998, pp. 383–398.

[Bell et al. 95]
John T. Bell and Folger Scott: "The investigation and Application of Virtual Reality as an Educational Tool," Proceedings of the American Society for Engineering Education 1995 Annual Conference, June, 1995.

[Birch et al. 95]
M. Birch, C. M. Boroni, F. W. Goosey, S. D. Patton, D. K. Poole, C. M. Pratt and R. J. Ross: "DYNALAB: A Dynamic Computer Science Infrastructure Featuring Program Animation," in Proceedings of the SIGCSE Session, ACM Meetings March 1995, Nashville, Tennessee.

[Bloom 56]
B. S. Bloom (Ed.): Taxonomy of educational objectives: The classification of educational goals: Handbook I, cognitive domain. New York 1956; Toronto: Longmans, Green.

[Brown 87]
M. Brown: Algorithm Animation, Cambridge MA, MIT Press 1987.

[Cross/Angelo 88]
K. P. Cross and T. A. Angelo, Classroom Assessment Techniques: A Handbook for Faculty 1988,

[Felder 93]
Richard M. Felder: "Reaching the Second Tier — Learning and Teaching Styles in College Science Education," Journal of College Science Teaching, pp 286–290, 23 (5), March/April, 1993.

[Howard et al. 96]
R. Howard C. Carver and W. Lane: "Felder's Learning Styles, Bloom's Taxonomy, and the Kolb Learning Cycle: Tying It All Together in the CS2 Course," in Proceedings of the SIGCSE Session, ACM Meetings, Philadelphia, Pennsylvania, February, 1996, pp. 227–232.

[Kolb 84]
D. A. Kolb: Experiential Learning: Experience as the Source of Learning and Development, Prentice-Hall, 1984.

[Kumar 00]
Amruth Kumar: "Dynamically Generating Problems on Static Scope," in Proceedings of the 5th SIGCSE/SIGCUE Conference on Innovation and Technology in Computer Science Education (Helsinki, Finland, July, 2000).

[McCarthy 87]
B. McCarthy: The 4MAT System. Teaching to Learning Styles with Right/Left Mode Techniques, EXCEL, 1987.

[Naps et al. 00]
T. Naps, J. Eagan and L. Norton: "JHAVÉ – An Environment to Actively Engage Students in Web-based Algorithm Visualizations", in Proceedings of the SIGCSE Session, ACM Meetings March 2000, Austin, Texas.

[Petre et al. 98]
M. Petre, A. Blackwell T. Green: "Cognitive Questions in Software Visualization," in Software Visualization: Programming as a Multimedia Experience, The MIT Press, 1998, pp. 453–480.

[Roessling et al. 00]
Guido Roessling, Markus Schueler and Bernd Freisleben: "The Animal Algorithm Animation Tool," Proceedings of the 5th Annual SIGCSE/SIGCUE Conference On Innovation and Technology in Computer Science Education, July, 2000, Helsinki, Finland.

[Stasko/Lawrence 98]
J. Stasko and A. Lawrence: "Empirically Assessing Algorithm Animations as Learning Aids," in Software Visualization: Programming as a Multimedia Experience, The MIT Press 1998, pp. 419–438.

[Stasko 90]
John Stasko: "Tango: A Framework and System for Algorithm Animation" in IEEE Computer, September, 1990.

[Stern et al. 99]
L. Stern, H. Sandergaard and L. Naish: "A Strategy for Managing Content Complexity in Algorithm Animation", Proceedings of ITiCSE'99, pp. 127-130, 1999.

User Interface Generation: Current Trends

María Dolores Lozano, Pascual González and Isidro Ramos

Various studies have proved that 48% of application code is devoted to the user interface and that nearly 50% of the implementation time is devoted to implementing it. The more easy the interface is to use, the more difficult is its development. In fact, the user interface is becoming more and more important as a means of interacting with a system. For this reason, in this article we have tackled the subject of user interface generation within the research area of Human-Computer Interaction.

Keywords: User Interfaces, Object Oriented Specification and Modelling Techniques

1 Introduction

New technologies and the user's constant demand for information have made the design and development of user interfaces more and more important in the area of software engineering. In fact, all aspects of Human-Computer Interaction (HCI) are being considered as a research field of great relevance and perhaps the one which has the greatest impact on the final users of information systems. Several disciplines can be considered within this research area, which range from the definition of user interface analysis and design methods, through the creation of new methods of representing information, to the study of the human factors involved in the management of such interfaces. In this article we review the most interesting investigations made up to now in the field of user interface generation and definition methods. Finally a new proposal is suggested, a user interface development environment within the software development process conforming to the object oriented model proposed in OASIS [Letelier et al. 98].

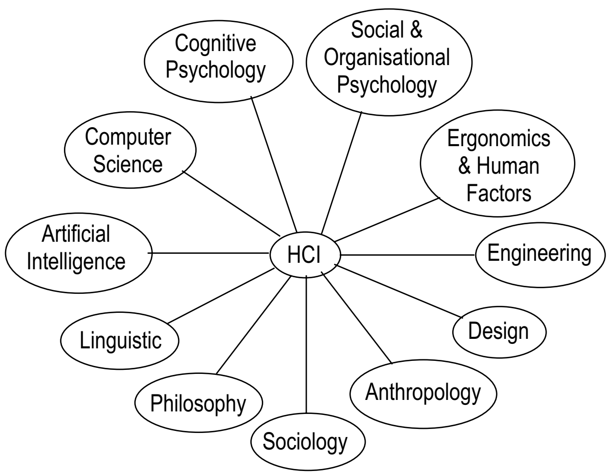


Fig. 1: HCI components.

2 Human Computer Interaction (HCI).

HCI is a discipline covering the design, evaluation and implementation of interactive systems for human use and entails the study of all the phenomena around them [ACM 92]. In the middle of the 1980's, the term HCI was adopted to reflect every factor involved in the interaction between users and computers.

As a matter of fact this term is becoming more and more important, mainly due to the "invasion" of computers in people's everyday activities. The computing evolution has contributed to this and has led to a greater preoccupation with user interface design when acknowledging that computer

María Dolores Lozano received her M.Sc. degree in Computer Engineering from the Technical University of Valencia. She is currently working on her Ph.D. degree in the area of user interface generation. In 1997 she joined the Department of Computer Engineering at the University of Castilla-La Mancha where she is an Assistant Professor for Software Engineering. Her research interests include human-computer interaction, user interface development, modelling techniques, object-orientation and specification languages. <mlozano@info-ab.uclm.es>

Pascual Gonzalez received the PhD. degree in Computer Engineering from the Technical University of Madrid. In 1991 he joined the Department of Computer Engineering at the University of Castilla-La Mancha where he is an Associate Professor for Software Engineering and director of the Information Technology Department in the Regional Development Institute. His research interests include human-computer interaction, user interface design, 3D interaction techniques, patterns, modeling techniques. <pgonzales@info-ab.uclm.es>

Isidro Ramos Salavert, is Full Professor in the Computation and Information Systems Dept. of the Technical University of Valencia, with 31 years of experience in different universities (Complutense de Madrid, Politécnica de Valencia, del País Vasco, Nancy). His research interests are: Object Oriented Conceptual Modelling, Foundations of Object Orientation. He has been a key researcher in 19 Research Projects, a member and president of the Programme Committee of relevant national and international conferences (CAISE, DEXA, EDBT…), and author of about 40 papers. He is presently responsible for International Coordination of the VII CYTED Subprogram on Applied Electronics and Computing. <iramos@dsic.upv.es>

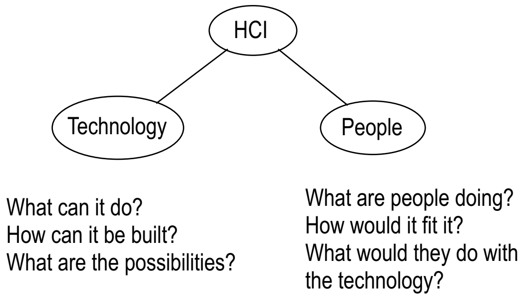


Fig. 2: HCI: Technological-Human Factors

programs should be designed to cover the user's needs and capabilities. That is, the user should not need to make radical changes to adapt himself to the system but the system must be designed to adjust to the user's needs.

Underlying HCI research is the belief that the people who use computer systems are the most important element. Their needs, capabilities and preferences to carry out their activities should dictate the design, implementation and use of the systems.

Human-Computer Interaction is a multidisciplinary science that not only includes computing or technological issues but also takes into account the other disciplines depicted in fig. 1.

The impact of engineering and design disciplines in HCI has been through Software Engineering due mainly to the fact that analysis and design methods, above all from the graphical point of view, have influenced the design of user interfaces.

The challenge of HCI is to acquire the maximum knowledge of technological capability and users needs so as to develop new technological artifacts that help and facilitate the human-machine interaction as depicted in figure 2.

3 User Interface Development Environments. State of the Art.

In this section we briefly describe the most outstanding features of current model-based user interface development environments. In a state of the art report, B. Myers presents a classification of user interface software tools [Myers 95]. This classification is based on the way that user interface developers can specify the layout and the dynamic behaviour of a user interface. This classification is as follows:

- *Language-based Tools:* These tools require the developer to program in a special purpose language.
- *Interactive Graphical Specification Tools:* These tools allow an interactive design of the user interface.
- *Model-based Generation Tools:* These tools use a high-level specification or model to generate the user interface automatically.

The last classification seems to be the most suitable to tackle user interface development. The first two support the specification of either the dynamic behaviour or the layout of the user interface in an easy way, but not both parts at the same time. In the same was as most current tools, they only support the last phase of the user interface life-cycle development, and the abstractions they provide do not have a direct connection with the task analysis results.

The typical components and the development process within a model-based user interface development environment are shown in figure 3.

The main component of these environments is the User Interface Model, which includes different declarative models such as the task, dialogue, domain, user and presentation models. Each one of them represents the different kinds of information needed for the final construction of the interface. Interactive and automatic development tools also have a role in this environment. The automatic generation tools are responsible for the transformations between the different declarative models and the final implementation of the desired user interface. The information required by the automatic generation and design tools is contained in the knowledge base.

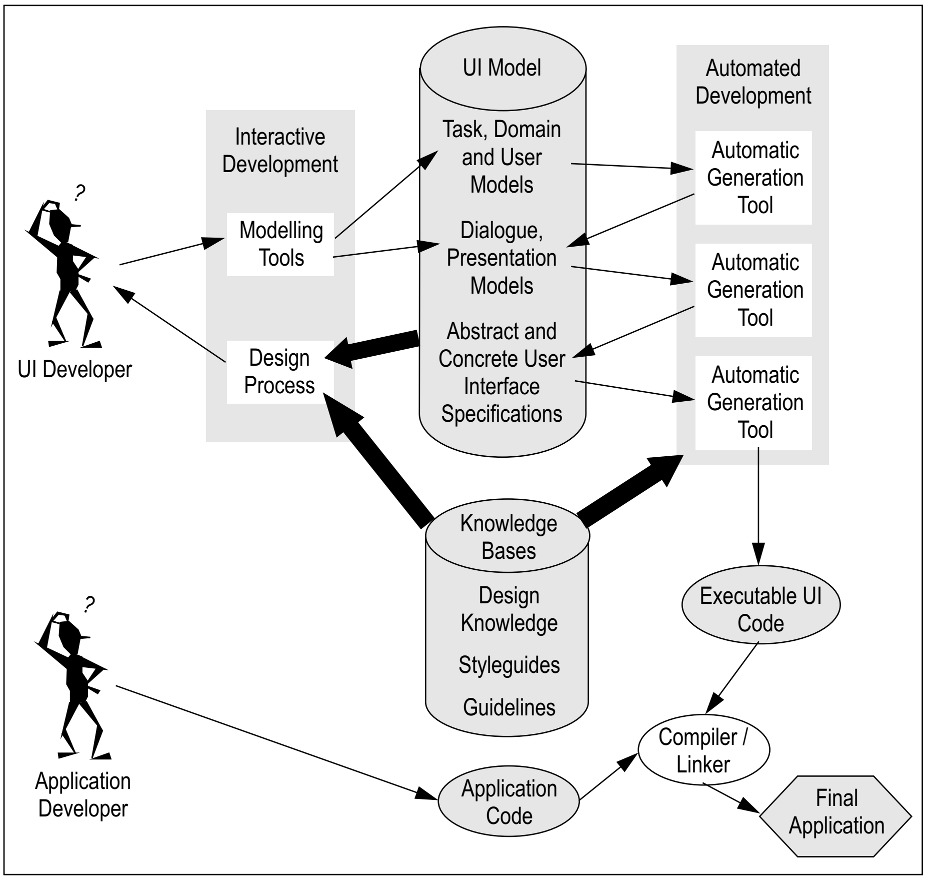


Fig. 3: Generic Architecture of a Model-based User Interface Development Environment.

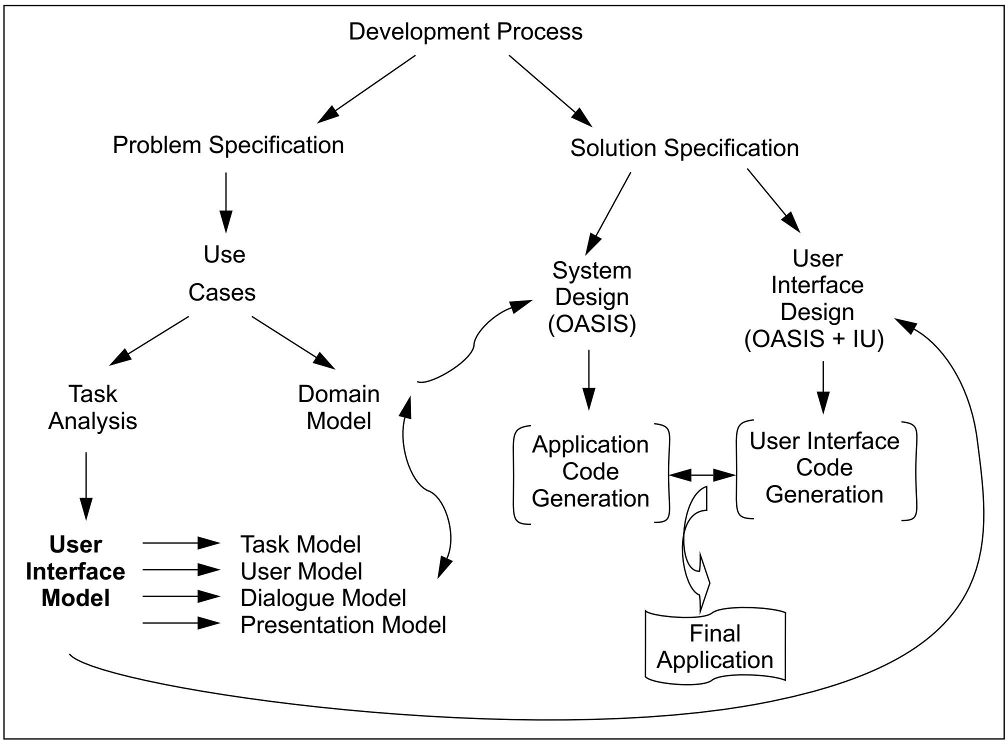


Fig. 4: IDEAS: Interface Development Environment within OASIS.

4 IDEAS: User Interface Generation within an Object Oriented Software Development Environment.

In this section we propose the integration of a user interface model within the software development environment defined by the object oriented model OASIS [Letelier et al. 98] from which the interface of an application can be generated. The user interface specification process is tackled in parallel to the application development. This development process is depicted in figure 4.

IDEAS [Lozano/Ramos 99b], [Lozano et al. 00] aims to be a model-based and automatic user interface development system integrated within the framework of OASIS to support the automatic production of high-quality user interfaces. This is possible due to the fact that this environment is based on declarative models, with the advantages of this approach as stated in the previous section.

Next, we describe this development process by levels and following the order in which the different models of our proposal are performed. This development process is depicted in figure 5.

At requirements level, the use cases [Jacobson et al. 92] technique is used. In this first stage and taking into account the use cases, the different kinds of users are identified. Subsequently, each use case is refined and the business rules, entities and actors that participate in it are specified.

At analysis level three models are generated. Firstly the Task Model, identifying the task the user has to perform with the system. This is done starting from the previous definition of use cases.

Following this, the Domain Model is performed. This model consists of two diagrams. The first one is the Sequence Diagram, which defines the system behaviour. The second one is the Roles Model, which defines the structure of the classes that take part in the associated sequence diagram together with the relationships among these classes, specifying the role of each one of them. Once the task and domain models have been defined, the user model can be defined. For every kind of user, the subset of tasks the user is allowed to perform is established. And for every task, a projection of the actions within the task that the user is allowed to accomplish is established. Finally, and according to the user's particular characteristics, the most appropriate way to show the information is defined.

At design level, the Dialogue Model is performed [Lozano/Ramos 99a]. All the models that have been generated up to now do not contain any graphical aspect of the final user interface. It is from now on that these issues start to be addressed and the way in which the user-system interaction will be performed is especially important. The dialogue model includes the generation of two different kinds of diagrams, the Dialogue Structure Diagram and the Component Specification Diagram. The first diagram specifies the user interface behaviour, this is, it represents the windows and dialogues that the user needs to complete all the tasks he requires from the system, and the user selections to pass from one window to another. For the generation of this diagram, we have to take into account the sequence diagram obtained in the analysis phase. An example of this type of diagram is shown in figure 6.

The second diagram entails establishing, for every one of the windows and dialogues obtained in the first diagram, the set of

to Contents, page 1

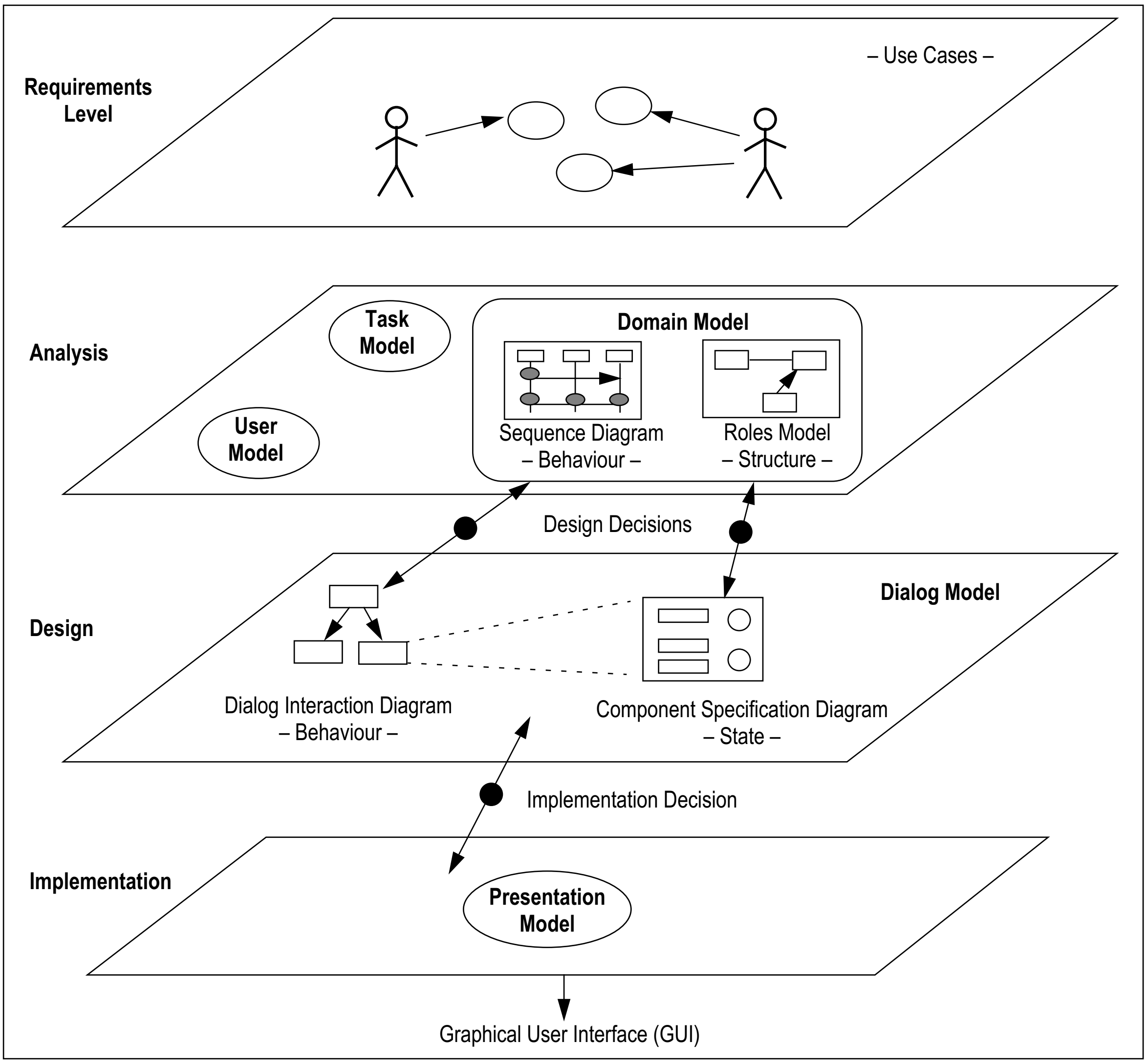


Fig. 5: User Interface Development Process.

control and visualization tools needed to give the user the functionality he requires to perform the corresponding task. This diagram defines the state of the user interface as depicted in figure 7:

At implementation level the Presentation Model is performed. In this model, the actual interaction objects comprising the Graphical User Interface are defined taking into account implementation decisions, the final implementation platform and according to the style guide followed.

5 Conclusions

The user interface is becoming more and more important as a means of interacting with a system. For this reason, in this article we have tackled the subject of user interface generation within the research area of Human-Computer Interaction.

With the aim of achieving an object oriented environment for automatic software production, that not only can be used with the efficiency and simplicity that the end-user requires through a friendly but at the same time robust interface, we have presented IDEAS, a user interface development environment with the philosophy of OASIS.

IDEAS, based on declarative models, provides descriptive information of the future components of the interface at a high level of abstraction. In the early development phases it also allows the study of how the interaction between the user and the system is performed. In the last phases it allows the automatic code generation based on the models.

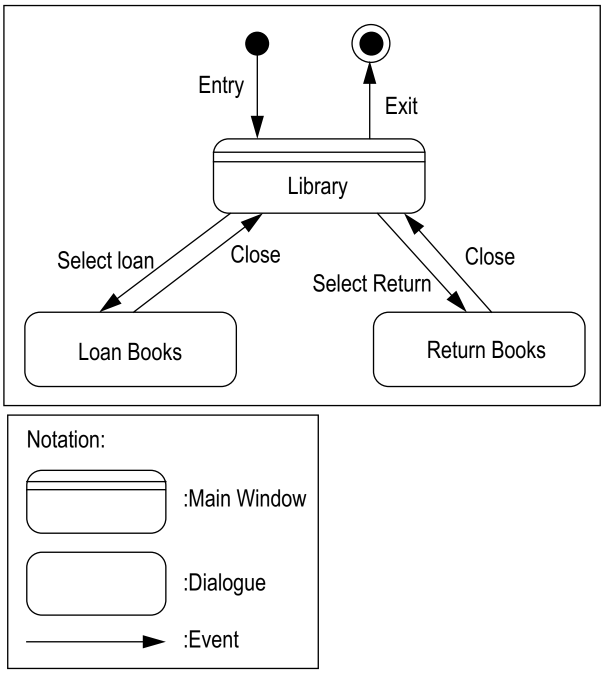


Fig. 6: Dialogue Interaction Diagram.

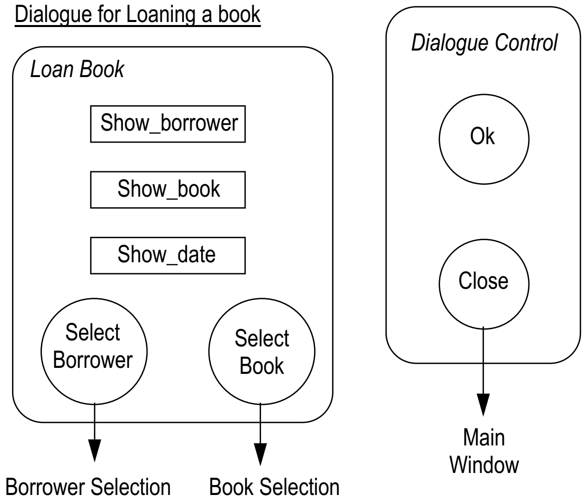


Fig. 7: Component Specification Diagram.

References

[ACM 92]
ACM Special Interest Group on Computer-Human Interaction. Curricula for Human-Computer Interaction. 1992.

[Jacobson et al. 92]
I.Jacobson, M. Christerson, et al.: Object-Oriented Software Engineering: A Use Case Driven Approach. Addison-Wesley, 1992.

[Letelier et al. 98]
P. Letelier, I. Ramos, P. Sanchez, O. Pastor. OASIS version 3.0: A Formal Approach for Object Oriented Conceptual Modelling. SPUPV-98.4011. Edited by the Technical University of Valencia, Spain, 1998.

[Lozano/Ramos 99a]
M.Lozano, I. Ramos: An Object Oriented Approach for the Specification of User Interfaces. PACRIM'99, Victoria -Canad, August, 1999.

[Lozano/Ramos 99b]
M. Lozano, I. Ramos: Integration of the User Model and Human Factors in an Object Oriented Software Production Environment. ECOOP'99, Lisboa -Portugal, June, 1999.

[Lozano et al. 00]
M.Lozano, I. Ramos, P. González. User Interface Specification and Modeling in an Object Oriented Environment for Automatic Software Development. 34th International Conference on Technology of Objetct-Oriented Languages and Systems. TOOLS-USA 2000. Santa Barbara, CA. 30 July – 3 August, 2000

[Myers 95]
B. A. Myers: User Interface Software Tools. ACM Transactions on Computer-Human Interaction 2 (1995), 1, 64–103

Visualization Designs for Constraint Logic Programming

Manuel Carro and Manuel Hermenegildo

We address the design and implementation of visual paradigms for observing the execution of constraint logic programs, aiming at debugging, tuning and optimization, and teaching. We focus on the display of data in CLP executions, where representation for constrained variables and for the constrains themselves are seeked. Two tools, VIFID and TRIFID, exemplifying the devised depictions, have been implemented, and are used to showcase the usefulness of the visualizations developed.

Keywords: Logic Programming, Constraint Logic Programming, Visualization, Debugging, Performance, Abstraction of Visual Representation.

1 Introduction

Program visualization has been classically used by computer scientists for many different purposes, including teaching, debugging, and optimization. However, classical program visualizations are often too dependent on the programming paradigms they were devised for, and do not adapt well to the nature of the computations performed in other paradigms (e.g., visualization of concurrent programs focuses on aspects which are not present in sequential programming). In particular, Constraint Programming nature differs radically from that of other programming paradigms and the visualization should address different problems. Moreover, it appears that the needs of CP practitioners are also different from those using other paradigms.

In any case, a good pictorial representation is fundamental to achieve a useful visualization. Thus, it is important to devise depictions which are well suited to the characteristics of CLP data and control. In addition, a recurring problem in the graphical representations of even medium-sized executions is the huge amount of information that is usually available. To cope successfully with these undoubtedly relevant cases, *abstractions* of the representations are needed. Ideally, such abstractions should show the most interesting characteristics (according to the particular objectives of the visualization process, which may be different in each case), without cluttering the display with unneeded details.

2 Constraint Logic Programming in a Nutshell

Constraint Programming (CP) is "one of the most exciting developments in programming languages of the last decade" [Marriot/Stuckey 98]. CP refers to programming using the equations which characterize the solution to a problem. These equations, which are similar to the arithmetical ones, can range over a wide gamut of domains: integers, reals, terms (i.e., data structures), strings, sets,… A CP system would automatically (and incrementally) solve these equations, therefore yielding a solution to the initial problem. This approach has undoubtedly much to do with, on one hand, mathematics itself, and, on the other hand, Operation Research. But, from a practical point of view, it departs from them in two ways: the ability to set up dynamically the equations which model a problem (and probably retract some of them at some point and add new ones), and the use of domains not usually found in mathematics.

Constraint Logic Programming (CLP) [Van Hentenryck 89] merges the constraint-based approach with the Logic Programming (LP) ideas, resulting in a highly synergetic combination. The properties of logic programming variables (single assignment, unification) and the control usually implemented in LP (automatic search procedures with backtracking, goal delaying) fit particularly well within constraint programming. The result is a family of languages which naturally extends LP in a unified framework (such that LP can be seen as a case of the

Manuel Hermenegildo is Full Professor at the Dept. of Artificial Intelligence in the School of Computer Science of the Technical University of Madrid, where he leads the Computational Logic, Implementation, and Parallelism (CLIP) group. He received his Masters and Ph.D. degrees in Computer Engineering from the University of Texas at Austin. He has participated in several EU funded (ESPRIT) and locally funded research projects as research group leader or coordinato. Previously he was project leader at the Microelectronics and Computer Technology Corporation (MCC) research centre and Adjunct Professor at the Department of Computer Science of the University of Texas at Austin. His research interests include advanced compilation techniques, parallelizing compilers, high-performance and distributed computing, concurrent and constraint logic programming, program execution visualization. He is editor of the Journal of Logic Programming. <herme@fi.upm.es>

Manuel Carro is currently assistant professor at the department of Languages and Systems, finishing his Ph.D. degree under the supervision of Manuel Hermenegildo. He received his Masters degree in Computer Science from the School of Computer Science of the Technical University of Madrid. His current research interests include abstract machines for parallel execution, visualization of constraint logic programming, implementation techniques and object oriented and distributed computing for logic and constraint logic languages. <mcarro@ffi.upm.es>

Fig. 1: Depiction of a FD variable

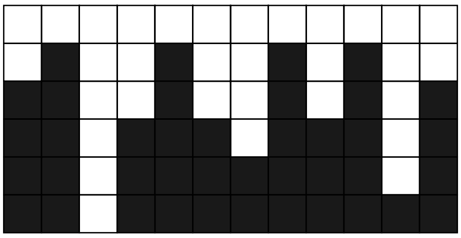


Fig. 2: History of a single variable

more general CLP), patching up some weaknesses (especially in arithmetic) found in LP languages.

One of the most useful CLP domains is finite domains: finite domain variables range over finite sets of integers, and the operations allowed between them are pointwise extensions of the regular arithmetic operations and relations. Although there is no elimination procedure similar to that of linear equations over the reals, a set of equations in finite domains can always be decided to have or not solution, ultimately thanks to the finiteness of the domains, using a mixture of simplification, value propagation, and labelling (i.e., assignment of values to variables). We will focus mainly, due to their practical importance, in finite domain variables.

3 Displaying Constrained Variables

The concept of variable binding in CLP is more complex than in imperative and functional languages: the value of a CLP variable is actually a complex object representing a (potentially infinite) set of values plus the constraints attached to the variable which relate it with the rest of the variables. Textual representations are usually not very informative and difficult to interpret and understand, and a graphical depiction of the variables can offer a view that is easier to grasp. Also, if we wish to follow the history of the program it is desirable that the graphical representation be either animated or laid out spatially as a series of pictures. The latter allows comparing different behaviours easily, trading time for space.

```
:- use_module(library(clpfd)).
:- use_module(library(tracing_library)).
program:-

        Variables = .... ,
        Names = .... ,
        open_log(Variables, Names, Handle),          %% Added
        constrain_values(Variables, Handle),
        log_state(Handle)                            %% Added
        visual_labelling(Variables, Handle),
        close_log(Handle).                           %%Added
```

Fig. 3: an annotated program skeleton

```
visual_labelling([],_Handle).
visual:labelling ([Q|Qs], Handle):-
        labelling([Q]),
        log_state(Handle),
        visual_labelling(Qs, Han-
```

Fig. 4: The `visual_labelling/2` library predicate

Depicting Finite Domain Variables

FD variables are instantiated to an initial domain, which is narrowed as equations are incrementally added and as the constraint system is simplified (either by algebraic rewriting or by the labelling procedure). At any state in the execution, each FD variable has an active domain (the set of allowed values for it) which is usually accessible by means of language primitives. For several reasons (space limitations, speed of addition/removal of constraints, etc.) this domain is usually represented using an upper approximation of the actual set of values that the variable can take.

A possible graphical representation is to assign a dot (or, depending on the visualization desired, a square) to every possible value a variable can take, highlighting those values in the current domain. An example of the representation of a variable `X` with current domain {1,2,4,5} from an initial domain {1…6} is shown in Figure 1.

It is extremely interesting to follow the evolution of a set of program variables throughout the execution. Probably the most useful portray is to simply stack the different state representations, as in Figure 2. It can reflect time accurately (for example, by mapping it to the height between changes) or ignore it by simply stacking a new row of a constant height every time a variable domain changes or an enumeration step is performed. This representation allows the user to perform an easy comparison between states and has the additional advantage of allowing more time-related information to be added to the display. Other possibilities we will not explore include animating the display, so that time is represented as such, or using different colour hues or shades of grey.

The *stacking* approach is one of the visualizations available in *VIFID*. *VIFID* is a Prolog library which represents the state of variables as instructed by spy-points introduced by the user in the program. Figure 3 shows an skeleton example of such an annotated program. The `open_log/3` primitive initializes the Handle data structure which contains the Variables to be observed and their Names. `close_log/1` takes the necessary actions in order to finish the visualization (e.g., closing a file, sending appropriate messages to the windows, etc.). The actual step-by-step depiction of the current state is made by the `log_state/1` primitive. It contacts with the visual side of the tool in order to communicate the current state of the var-

to Contents, page 1

Fig. 5: Evolution of FD variables for a 10-queens problem

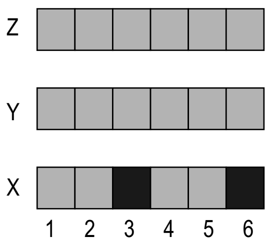


Fig. 6: Several variables side to side

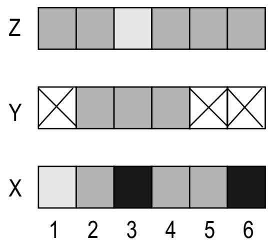


Fig. 7: Changing a domain

iables and update the windows.

An important part of the CLP execution is the labelling phase, which tries to assign values to the variables which are compatible with the existing constraints. This labelling is usually performed by a built-in, which receives a list of variables and an indication of the labelling strategy. We visualize the evolution of the variables during labelling by recoding this built-in so that the state of the variables is logged after each labelling step. Figure 4 shows an example implementation, which receives the list of variables to label and the Handle to the visualization and performs a tailored labelling. It is a simplified code for illustration purposes, but it clarifies how this (and other primitives) can be interfaced with the visual tools without too much effort.

Figure 5 shows a screen dump of a window generated by *VIFID* presenting the evolution the state variables when solving the Queens problem for a board of size 10. Each column in the display corresponds to one program variable. The possible values are the row numbers in which a queen can be placed. Lighter squares represent values still in the domain, and darker squares represent discarded values. Each row in the display corresponds to a spy-point in the source program, and points where backtracking happened are marked with small hooks. It is straightforward to see that very little backtracking was necessary, and that variables are highly constrained, so that enumeration (proceeding left to right) quite quickly discarded initial values.

4 Representing Constraints

It is obviously interesting to represent the relationships among several variables as imposed by the constraints affecting them. Textual representation is often not straightforward (or even possible in some constraint domains), can be computationally expensive, and provides too much level of detail for an intuitive understanding. Moreover, in general there are many states of the variables which meet the restrictions imposed by the constraints. A general solution which takes advantage of the representation of the actual values of a variable (and which is independent of how this representation is actually performed) is to use projections to present the data piecemeal and to allow the user to update the values of the projected variables, while observing how the variables being shown are affected by such changes. This can often give the user an intuition of the relationships linking the variables (and detect, for example, the presence of erroneous constraints). We will use the constraint **C1**, below, in the examples which follows:

$$\mathbf{C1} \equiv X \in \{1..6\} \wedge X \neq 6 \wedge X \neq 3 \wedge Z \in \{1..6\} \wedge Z = 2X - Y \wedge Y \in \{1..6\}$$

Figure 6 shows the domains of FD variables `X`, `Y`, and `Z` subject to **C1**. An update of the domain of a variable should induce changes in the domains of other related variables. For example,

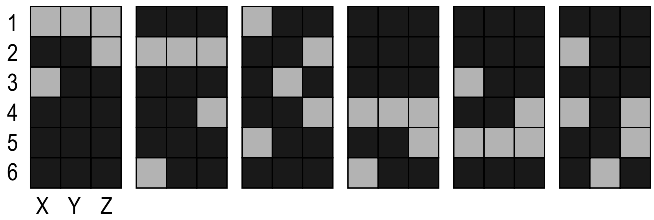


Fig. 8: Enumerating `Y`, representing enumerated domains for `X` and `Z`

we may discard the values 1, 5, and 6 from the domain of `Y`, which boils down to representing the constraint $\mathbf{C2} \equiv \mathbf{C1} \wedge Y \neq 1 \wedge Y < 5$

Figure 7 represents the new domains of the variables. Values directly disallowed by **C2** are shown as crossed boxes; values discarded by the effect of this constraint are shown in a lighter shade. In this example the domains of both X and Z are affected by this change, and so they depend on Y; this type of user-driven visualization is also available in the *VIFID* tool. A more detailed inspection can be done by leaving just one element in the domain of a variable, and watching how the domains of other variables are updated. In Figure 8 `Y` is given a definite value from 1 (in the leftmost rectangle) to 6 (in the rightmost one). This allows the programmer to check that simple constraints hold among variables, or that more complex properties (e.g., that a variable is made definite by the definiteness of another one) are met.

A static version of this view can be obtained by plotting values of pairs of variables in a 2-D grid. This is schematically shown in Figures 9, 10, and 11, where the variables are subject to the constraint **C2**. From these representations we can deduce that the values `X = 3` and `X = 6` are not feasible, regardless of the values of `Y` and `Z`. It turns out also that the plots of `X` against `Y` and `X` against `Z` (Figures 9 and 10) are identical. From this, one might guess that perhaps `Y` and `Z` have necessarily the same value, i.e., that the constraint `Z=Y` is entailed by the store. This possibility is discarded by Figure 11, in which we see that there are consistent pairs where `X≠Z`. Furthermore, the slope of the highlighted squares on the grid suggests that there is an inverse relationship between `Z` and `Y`: incrementing one of them would presumably decrement the other – and this is actually the case, from constraint **C1**. A *VIFID* window showing a 10-Queen 2-D plot appears in Figure 12; the check buttons at the bottom allow the user to select the variables to depict.

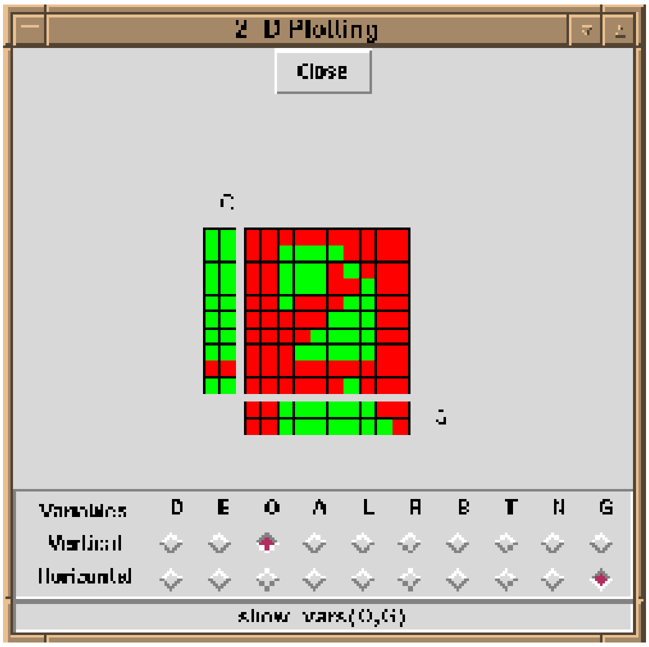


Fig. 12: Relating variables in *VIFID*

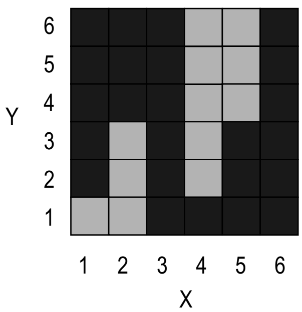


Fig. 9: `X` against `Y`

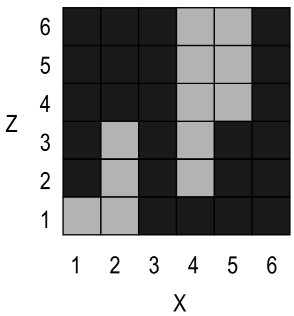


Fig. 10: `X` against `Z`

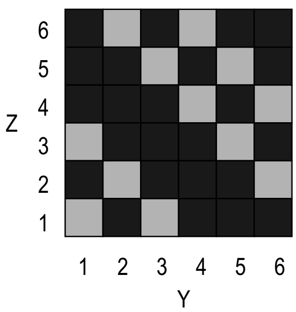


Fig. 11: `Y` against `Z`

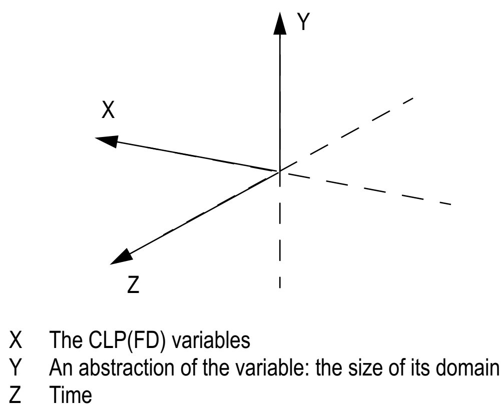


Fig. 13: Meaning of the dimensions in the 3-D representation

```
:- use_module(library(clpfd)).
:- use_module(library(trifid)).

dgr(ListofVars):-
    ListOfVars = [G,O,B,N,E,A,R,L,T,D],
    open_log(ListOfVars, Handle),                   %% Added
    domain(ListOfVars, 0, 9),
    log_state(Handle),                              %% Added
    D #> 0,
    log_state(Handle),                              %% Added
    G #> 0,
    log_state(Handle),                              %% Added
    all_different(ListOfVars),
    log_state(Handle),                              %% Added
    100000*D + 10000*O + 1000*N + 100*A + 10*L + D +
    100000*G + 10000*E + 1000*R + 100*A + 10*L + D #=
    100000*R + 10000*O + 1000*B + 100*E + 10*R + T,
    log_state(Handle),                              %% Added
    visual_labelling(ListOfVars, Handle),
    close_log(Handle).
```

Fig. 14: The annotated `DONALD + GERALD = ROBERT` FD program

5 Abstraction for Constraint Visualization

It is often the case that executions of large programs result in too much data being displayed. Even if an easy-to-understand depiction is provided, the amount of data can overwhelm the user with an unwanted level of detail. Abstraction is a method to cope with this problem.

Abstracting Values

While the presence of a large number of variables can be partially solved by a careful selection of variables, another problem remains: representations of variables with a large number of possible values can convey information too detailed. At the limit, the screen resolution may be insufficient to assign a pixel to every value in the domain. This is easily solved by using scrollable canvases, providing means for zooming, fish-eye views, etc. That was the approach taken in the VisAndOr tool [Carro et al. 93] aimed at showing the parallel execution of logic languages.

However, these methods are more "physical" approaches than true conceptual abstractions of the information, which are richer and more flexible.

An alternative is to use an application-oriented filtering of the variable domains. For example, if some parts of the program are trusted, their effects can be masked out by removing the values already discarded from the representation of the variables: e.g., if a variable is known to take only odd values, the even values are simply not shown in the representation. This filtering can be specified using the source language – in fact, the constraint which is to be abstracted should be the filter of the domain of the displayed variables.

Another alternative is to perform a more semantic "compaction" of parts of the domain. As an example, consider presenting the domain of a variable simply as a number, denoting how many values remain in its current domain, thus providing an indication of its "degree of freedom". This idea is the basis of our next visualization.

Domain Compaction and New Dimensions

Besides the problems in applications with large domains, the static representations of the history of the execution (Figure 5) can also fall short in showing intuitively how variables converge towards their final values, again because of the excess of points in the domains, or because an execution shows a "chaotic" profile. A better option is to use the number of active values in the domain as coordinates in an additional dimension. Figure 14 shows a CLP(FD) program for the `DONALD+GERALD=ROBERT` puzzle. The program was annotated with calls to predicates which act as spy-points, and log the sizes of the domains of each variable at the time of each call.

Figure 15 is an execution of the program in Figure 14. The variables closer to the origin (the ones which were labelled first) are assigned values quite soon in the execution and they remain fixed. There are backtracking points scattered along the execution, which appear as blocks of variables protruding out of the picture. There is also a variable (viewed as a white strip in the middle of the picture) which appears to be highly constrained, so that its domain is reduced right from the beginning; that variable is probably a good candidate to be labelled soon in the execution. Some other variables apparently have a high interdependence (at least, from the point of view of the solver): in case of backtracking, the change of one of them affects the others. This suggests that the variables in this program can be classified into two categories: one with highly related variables (those whose domains change at once in case of backtracking) and a second one which contains variables relatively independent from those in the first set.

These figures have been generated by a tool, *TRIFID*, integrated into the *VIFID* environment. They were produced by

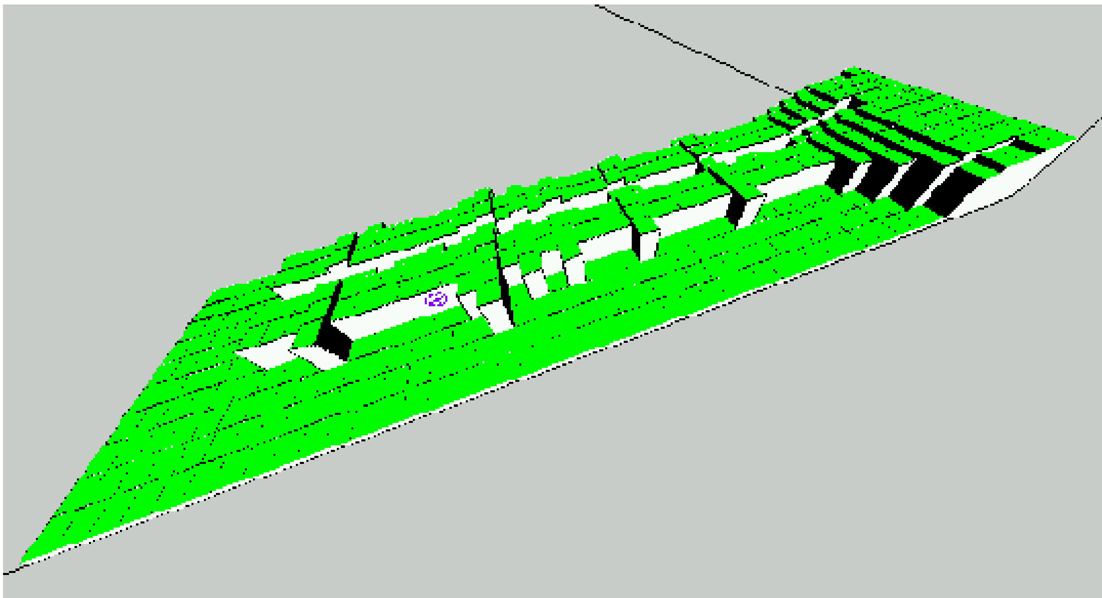


Fig. 15: Execution of the `DONALD + GERALD = ROBERT` program

using the ProVRML package [Smedbäck et al. 99], which allows reading and writing VRML code from Prolog, with a similar approach to the one used by PiLLoW [Cabeza/Hermenegildo 97]. One advantage of using VRML is that sophisticated VRML viewers are readily available for most platforms. The resulting VRML file can be loaded into such a viewer and rotated, zoomed in and out, etc. Another reason to use VRML is the possibility of using hyper-references to add information and animation to the depiction of the execution without cluttering the display.

Abstracting Constraints

As the number and complexity of constraints in programs grow visualizing them as relationships among variables may cause the same problems we faced when trying to represent values of variables. The solutions suggested for the case of representation of values are still valid and can give an intuition of how a given variable relates to others. However, it is not always easy to deduce from them how variables are related to each other, due to the lack of accuracy in their representation.

A different approach to abstracting the constraints in the store is to show them as a graph [Montanari/Rossi 91] where variables are represented as nodes which are linked if the corresponding variables are related by a constraint (Figure 16)[1]. This representation provides the programmer with an approximate understanding of the constraints in the solver (but not exactly *which* constraints they are). Moreover, since different solvers behave in different ways, this can provide hints about better ways of setting up constraints for a given program and constraint solver. The topology of the graph can be used to decide whether a reorganization of the program is advantageous; for example, if there are subsets of nodes in the graph with a high degree of connectivity, but those subsets are loosely interconnected, it may be worth to set up the tightly connected sections and make a (partial) enumeration early, to favour local constraint propagation, and then link the different regions. Animation can reflect propagation and how variables acquire a definite value. In Figure 17 some variables became definite, and as a result some constraints them are not shown any more: this reflects the idea of a system being progressively simplified, and also visualizes how backtracking affects the constraint store. Further filtering can be accomplished by selecting which types of constraints are to be represented (e.g, represent only "greater than" constraints, or certain constraints flagged in the program through annotations).

1. This particular figure is only appropriate for binary relationships; constraints of higher arity would need hypergraphs.

6 Implementation Details

VIFID and *TRIFID* are implemented in Prolog and Tcl/Tk, and rely on a few primitives to open socket connections and to spawn and communicate with other processes (primarily for the Tcl/Tk part). *VIFID* is completely interactive, and since the library has direct access to the program variables, the user can update them on the fly. The execution can continue after updating, but the user does not have to commit to this update: a RESET button forces the program to backtrack to the point where the update was made. Only a few routines commonly used throughout the execution were written directly in Tcl/Tk. The flexibility of Tcl/Tk was enough, since most of the windows have a simple layout. The speed of Tcl/Tk was not much of a problem, except when the number of objects in the window became very large. Overall, the tool was strong enough to be used routinely, and the visualization was found to be useful and easy to understand.

TRIFID shares many ideas with *VIFID*: It is also a Prolog library which scans the variables it has access to, but instead of

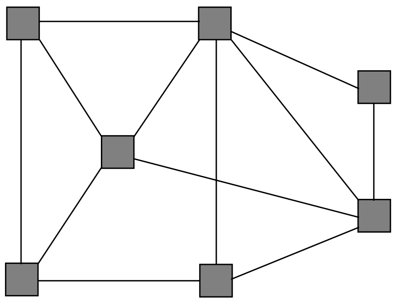


Fig. 16: Constraints represented as a graph

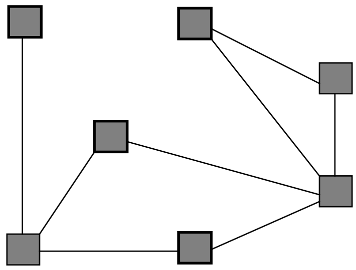


Fig. 17: Bold frames represent definite values

starting an interactive 3-D visualization, we decided to take advantage of a Prolog to *VRML* interface and generate *VRML*. Gathering the data was not a computational problem; instead, we found troubles related with the size of the generated files[2], and with the speed of the VRML visualizers (freely) available.

7 Related Work

Early work in constraint visualization was made for Eclipse [ECRC 93]; the GRACE system [Meier 96] represented the values of constrained variables as we did in Section 3.1, connected to a Byrd box model for program debugging. Additional information was encoded using different colour shades. More recently, the DisCiPl project [Deransart et al. 00] fostered the use of visualization and assertion-based debugging tools.

Some constraint applications need to set up complex relationships among the variables. In those cases a visualization which mimics the initial problem helps in mapping problems in the constraint solving to the original problem. The Global Constraint visualization tool [Simonis et al. 00] does precisely this, by incorporating special visualizations tailored to some of the complex constraints available in the CHIP system. Although this gives an intuitive representation, it needs the user to map the problem to one of these *standard* complex constraint templates.

The visualization of constraint networks, proposed here as an constraint abstraction amenable of being treated and studied, was implemented at a different level in the Constraint Investigator [Müller 99], interfaced with the Oz Explorer [Schulte 97]. This proposal visualizes a graph which is close to the implementation. The ability to expand and collapse the constraint net and to filter the variables increases the tool usefulness in the case of big executions. It gives a good representation of the store, but probably needs some further structure to represent complex problems.

8 Conclusions

We have discussed techniques for visualizing data evolution in CLP. The graphical representations have been chosen based on the perceived needs of a programmer trying to analyse the behaviour and characteristics of an execution. We have proposed solutions for the representation of the run-time values of the variables and of the run-time constraints among them. In order to be able to deal with large executions, we have also discussed some abstraction techniques, including the 3-D rendition of the evolution of the domain size of the variables. The proposed visualizations for variables and constraints have been tested using two integrated prototype tools: *VIFID* and *TRIFID*. *VIFID* and, to a lesser extent, *TRIFID*, which is less mature, have evolved into a practical system. Also, some of the views and ideas proposed have since made their way to other tools, such as those developed for the CHIP system [Simonis/Aggoun 00].

2. More precisely, the VRML visualizers had problems with that!

Acknowledgement

This work has been partially supported by Spanish EDIPIA Grant (CICYT TIC 99-1151).

References

[Carro et al. 93]
M. Carro, L. Gómez, and M. Hermenegildo. Some Paradigms for Visualizing Parallel Execution of Logic Programs. In 1993 International Conference on Logic Programming, pages 184–201. MIT Press, June 1993.

[Cabeza/Hermenegildo 97]
D. Cabeza and M. Hermenegildo. WWW Programming using Computational Logic Systems (and the PiLLoW/Ciao Library). In Proceedings of the Workshop on Logic Programming and the WWW at WWW6, San Francisco, CA, April 1997.

[Deransart et al. 00]
P. Deransart, M. Hermenegildo, and J. Maluszynski. Analysis and Visualization Tools for Constraint Programming. Number 1870 in LNCS. Springer-Verlag, September 2000.

[ECRC 93]
ECRC. Eclipse User's Guide. European Computer Research Centre, 1993.

[Meier 96]
M. Meier. Grace User Manual, 1996. Available at http://www.ecrc.de/eclipse/html/grace/grace.html.

[Montanari/Rossi 91]
U. Montanari and F. Rossi. True-concurrency in Concurrent Constraint Programming. In V. Saraswat and K. Ueda, editors, Proceedings of the 1991 International Symposium on Logic Programming, pages 694–716, San Diego, USA, 1991. The MIT Press.

[Marriot/Stuckey 98]
Kim Marriot and Peter Stuckey. Programming with Constraints: An Introduction. The MIT Press, 1998.

[Simonis/Aggoun 00]
H. Simonis and A. Aggoun. Search Tree Visualization. In P. Deransart, M. Hermenegildo, and J. Maluszynski, editors, Analysis and Visualization Tools for Constraint Programming, number 1870 in LNCS. Springer-Verlag, September 2000.

to Contents, page 1

[Simonis et al. 00]
H. Simonis, A. Aggoun, N. Beldiceanu, and E. Bourreau. Complex Constraint Abstraction: Global Constraint Visualization. In P. Deransart, M. Hermenegildo, and J. Maluszynski, editors, Analysis and Visualization Tools for Constraint Programming, number 1870 in LNCS. Springer-Verlag, September 2000.

[Schulte 97]
Christian Schulte. Oz explorer: A visual constraint programming tool. In Lee Naish, editor, ICLP'97. MIT Press, July 1997.

[Smedbäck et al. 99]
G. Smedbäck, M. Carro, and M. Hermenegildo. Interfacing Prolog and VRML and its Application to Constraint Visualization. In The Practical Application of Constraint Technologies and Logic programming, pages 453-471. The Practical Application Company, April 1999.

[Müller 99]
Tobias Müller. Practical Investigation of Constraints with Graph Views. Poster in International Conference on Logic Programming, December 1999.

[Van Hentenryck 89]
P. Van Hentenryck. Constraint Satisfaction in Logic Programming. MIT Press, 1989.

Software Visualization Generation Using ANIMALSCRIPT

Guido Rößling and Bernd Freisleben

Many different software visualization tools are available to the public. Most, however, only address specific application areas, forcing users to shift between tools. The visualization tool ANIMAL, with its built-in visualization language ANIMALSCRIPT, is an algorithm animation tool including a graphical editor that can be used for arbitrary software visualizations. This article introduces the basics of ANIMAL and ANIMALSCRIPT as well as some special features including internationalization support. ANIMALSCRIPT is also easy to adapt and extend with more advanced features.

Keywords: Software Visualization, Algorithm Animation, Internationalization Issues, Dynamic Placement, Dynamic Extension.

1 Introduction

The increasing popularity of algorithm or program visualization has led to the development of several tools, most of which are designed to address a certain field of visualization. The advantage of implementing such specific tools is that they may be highly specialized within their field and thus may offer great field-specific features. However, they are usually not usable outside their specific field. For example, a tool geared towards visualizing run-length encoding algorithms such as *RLE* [Khuri/Hsu 00] is highly useful in its field, but not usable for other topics. One of the areas that shows growing interest in visualizations is academic education. Here, a fair number of teachers are interested in using generated algorithm animations in their lectures instead of having to perform the operations themselves on the blackboard. However, if the tool used for the visualization is highly specialised, they will have to shift between using the tool and the normal mode of presentation. This becomes even more relevant in conferences, where presenters who want to show the usage of such a tool have to change between the visualization tool and the tool used for the presentation, such as *PowerPoint*™ .

Having a single tool that is generic enough to handle both the presentation and the content display saves the time needed to move between tools. More importantly, it also prevents the presenter from having to learn how to use a number of different tools. However, the tool must be sufficiently generic that it can handle the typical requirements of presentations as well as to support context-specific operations. It is unlikely that a single tool could offer built-in support for all the types of software visualizations that users might want. Therefore, it may be beneficial to have a tool that can be dynamically extended by users to address their specific needs without forcing them to read and modify large portions of the underlying code. This article presents the scripting language ANIMALSCRIPT that extends the capabilities of the ANIMAL visualization system [Rößling et al. 00]. ANIMALSCRIPT offers the basic operations required for software visualization, but can easily be adapted to address other context-specific requirements. For example, ANIMALSCRIPT supports relative object placement and addresses internationalization issues.

The article is organized as follows. Section 2 presents some tools specific to a certain usage area. Section 3 gives a short overview of the underlying visualization concept of ANIMAL and how visualizations can be generated. Section 4 presents the capabilities of ANIMALSCRIPT including several screen shots of the system. Some possibilities for adapting and extending the capabilities of both ANIMAL and ANIMALSCRIPT are outlined in section 5. Section 6 concludes the paper and outlines areas for further research.

2 Classification of Available Tools

Due to the increasing popularity of software visualization, many systems have been introduced during the last years. Since we cannot cover all available tools, we will focus on some tools that are representative of their categories. The tools can be divided into different categories based on the type of input required to generate a visualization: mouse actions within

Guido Rößling received the diploma degree in Computer Science in 1996 from the Darmstadt University of Technology. He has been a full-time research assistant at the University of Siegen since 1996 and from 1997, he has given lectures on Computer Science and Programming at the Berufsakademie Mannheim. He is currently working on his Ph.D. thesis which describes a dynamically extensible and configurable software visualization tool. His research interests include the effective use of multimedia in teaching, algorithm animation and automatic generation of data representations. He is a member of ACM, the IEEE Computer Society, AACE and GI. <roessling@acm.org>

Bernd Freisleben is associate professor of computer science in the department of Electrical Engineering and Computer Science at the University of Siegen. He received his Master's degree in Computer Science from the Pennsylvania State University, USA, in 1981, and the Ph.D. degree in Computer Science from Darmstadt University of Technology, in 1985. His research interests include distributed systems, network computing, and multimedia in education. <freisleb@informatik.uni-siegen.de>

to Contents, page 1

a GUI, source code, method calls in a specialized method library, or scripting language commands.

Classic presentation tools such as *PowerPoint*™ from Microsoft or Sun's *StarImpress*™ typically expect the user to provide all contents of the visualization manually within a graphical user interface (GUI). The advantage of this approach is that the user has complete control over all elements, at least within the scope of what the tool can do. However, being forced to do everything manually also means that the generation cannot be automated. If a second visualization of the same topic but with changed values is required, the user basically has to start again from scratch. Furthermore, presentation tools are often more geared towards generating stunning visual displays for short presentations than addressing the needs of presenting the semantics of dynamic processes such as algorithms or data structures. Thus, they often offer a fair selection of ways in which new elements can be added, but may be limited in how elements that are already visible may be modified. Furthermore, the underlying concept of the presentation is a sequence of slides. This means that the next slide will initially be empty and thus not reflect the contents of the last slide – which is unfortunate for visualizing algorithms that perform several operations spread over a number of "slides". While presentation tools are very useful for presentations, they are thus less useful for presenting the behaviour of systems or algorithms.

Source-code based visualization systems such as *DDD* [Zeller 00], *Jeliot* [Haajanen et al. 97] / *Jeliot 2000* [Ben-Ari 00] and *WinHIPE* [Naharro-Berrocal et al. 00] automatically present a visualization of the underlying source-code. This also means that it is very easy to automatically generate visualizations: the user only needs to provide the source code, or modify parameter values according to the specific requirements. However, a number of restrictions apply to these systems. Firstly, the programming language that can be used is fixed by the system. For example, *DDD* addresses machine-code programs whose input typically will come from *C* or *C++*, while *Jeliot* and *Jeliot 2000* work on Java source code and *WinHIPE* is restricted to functional programming. If the programming language used in the presentation context is not supported by the tool, the tool cannot be used. Due to the automatic generation of the visualization, the user will also have only a limited ability to change the display. For example, it is unlikely that the system will be able to deduce the semantically most appropriate representation of a given data structure from the underlying code, e.g. a stack implemented using a linked list might not be represented as a stack, but as a standard list. Finally, presenters might wish to add explanatory text, or take a more abstract view of the underlying algorithm (such as a pseudo-code implementation) to avoid confusing the viewers with implementation details. These types of tools do typically not support both operations. Finally, if the topic to be presented is not available as an algorithm, the tools cannot be used at all.

Tools such as Silicon Graphic's *JAL* [Silicon Graphics 98] rely on method calls within a specific programming library (API) for generating the visualization. The ease of use of the library, as well as the extent to which operations are supported, depends on the state of the library. Two effective limitations are clear: (1) the user has to be sufficiently proficient in programming to write a program that correctly invokes the library methods and (2) the language in which the library is implemented must match the language used for the underlying code. If both are true then visualization APIs can be very useful; otherwise, the library is inappropriate and thus useless for the target application.

Systems that expect input in a specific format such as a scripting language offer a partial solution to the problem of inappropriateness. The tool does not care how the scripting input is generated – be it manually by the user, or automatically by the underlying code. Thus, the user does not have to be proficient in programming. However, algorithms implemented in any language may be modified so that they generate the appropriate scripting commands. Due to this fact, tools that work on scripting input have become popular over the last few years: *JSamba* [Stasko 98], *JAWAA* [Pierson/Rodger 98] and highly specific tools such as *JellRep* and *JFlap* [Hung/Rodger 97] for compiler construction topics. The scripting language underlying *JSamba* is kept generic, while most other tools offer a scripting language geared towards a specific context. For example, *JAWAA* is geared for algorithm and data structures, thus offering elements such as trees and graphs. However, it is less useful for presentations, as it does not support elements such as item lists. Additionally, going backwards one or more steps in the visualization is often not supported by the tools.

Users should generally choose the type of tool that is most appropriate to their intentions. They should reflect on the characteristics and limitations of each approach, especially if they wish to reuse the tool in other, less specific applications. Scripting-based tools offer a reasonable compromise between ease of generation and automation support. If the scripting language is not highly specialized, they also offer easy reuse in other contexts.

In the following, we will examine the scripting language ANIMALSCRIPT. As ANIMALSCRIPT is embedded in a visualization system, we will first take a look at the central features of this system.

3 The ANIMAL Visualization System

The ANIMAL visualization system [Rößling et al. 00] treats a visualization as a sequence of discrete steps. Each step may contain an arbitrary number of transformations. Each transformation in turn works on an arbitrary number of objects at the same time. ANIMAL offers the basic graphic primitives point, polyline / polygon, arc and text. As a concession to its initial use for introductory computer science visualizations, it also offers list element primitives. Each of these primitives contains a number of user-adjustable properties. Some properties are common to all primitives, for example, colour and depth information to resolve display conflicts caused by overlapping elements. Other properties depend on the given primitive. For example, a polygon object may have a fill colour, while this has no semantic meaning for a polyline. On the other hand, polylines may have a pointer at the beginning or the end, which is not appropriate for polygons. Note that ANIMAL does not restrict the number of nodes a polyline can have.

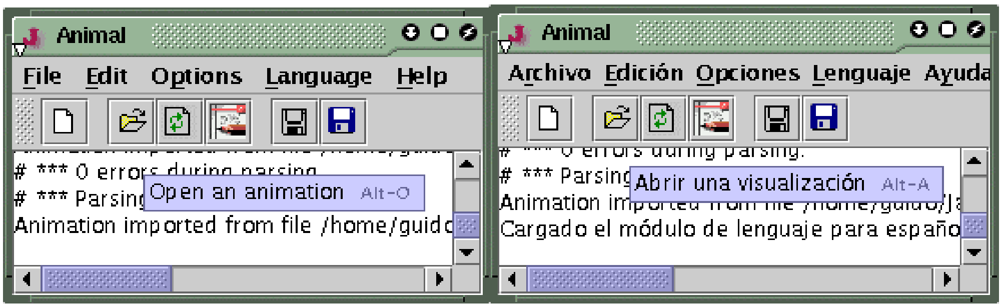


Fig. 1: How translating ANIMAL's GUI affects titles, messages and keyboard short-cuts

ANIMAL offers transformations for changing the visibility state of arbitrary objects, rotating or moving objects and changing their colour properties. The extent to which these operations are supported depends on the type of objects they are executed on. For example, when moving a list element, users can choose between moving the element as it is, adjusting only one or more list pointers, or moving the element with fixed pointers. An arbitrary polyline or arc can be used as the object along which objects are to be moved, enabling complex move operations. Colour changes are applicable to all colour properties of a given object, so it is possible to have a polygon with a fill colour different from the outline colour. All operations can be scheduled by specifying a start time and a duration.

The individual visualization steps are usually chosen by using the video player-like buttons provided. These also include functionality for going backwards in the visualization. The author of a visualization may also specify that a given step will automatically advance to the next step after a certain amount of time has passed. The user may use a slide ruler displaying the percentage of the visualization already shown to adjust the current state. Finally, the author may provide labels to steps that are collected in a separate window and which can be used to jump to the associated step from any place in the visualization with a single mouse click.

In order to make ANIMAL usable for the widest possible target audience, we have followed the principle "keep it simple". ANIMAL offers a small graphical interface allowing users to edit any presentation using the mouse with standard features such as a grid and drag and drop support. The number of both graphic primitives and transformations has been kept intentionally small. Users who want a more powerful interface are expected to move to the built-in scripting language ANIMALSCRIPT once they have become familiar with using ANIMAL. The graphical interface and all output messages can be translated into another language by a single mouse click. At the moment, we can only provide German and English language support. Support for other languages such as Spanish, French and Italian requires the translation of a text file, without recompilation. Figure 1 shows ANIMAL's main window after the language has been changed. Note that the translation affects menus and their items, explanatory texts and keyboard shortcuts.

ANIMAL also offers an interface for import filters and provides a number of export filters. Currently, visualizations can be imported only from *JSamba* but be exported as a QuickTime movie, a variety of image formats and in XML format. Other formats are relatively straightforward to implement if the precise definition of the target format is available. All filters are registered in a configuration file that can be updated whenever a new filter becomes available. Adding new import or export filters does not require access to the base ANIMAL code, as no recompilation is necessary.

4 Selected Features of ANIMALSCRIPT

ANIMALSCRIPT supports a large number of features, so that we cannot cover all of them here. Therefore, we will start with a short introduction to ANIMALSCRIPT and then focus on three special features. These are relative object placement, internationalization issues and diagnostic output support. ANIMALSCRIPT contains commands for all the features of ANIMAL listed in the previous section. It also includes specific commands for common subtypes, such as squares instead of only generic polygons. Like several other scripting languages for visualization [Pierson/Rodger 98], [Stasko 98], ANIMALSCRIPT is line-based – each line contains exactly one command or comment. All graphic elements used during a visualization have to be given a name to allow their use in later transformations. These names can be exchanged using a swap command, making many transformations much easier for the user.

Some of the more advanced features of ANIMALSCRIPT go beyond the operations available within the graphical user interface of ANIMAL. For example, a single scripting command is sufficient to link or unlink list elements. Other special commands built into ANIMALSCRIPT support array data structures including pointers to array cells, overwriting of cell values and swapping the contents of two cells in the same array. To provide greater flexibility for source or pseudo code usage, a special code environment is also provided that handles indentation and offers highlighting for the current line of code or the context, such as the enclosing loop statement.

The main difference between ANIMALSCRIPT and the scripting languages used in other systems such as *JAWAA* and *JSamba*, however, lies in the areas of relative element placement, internationalization and diagnostic output. The relative element placement of ANIMALSCRIPT frees authors from keeping track

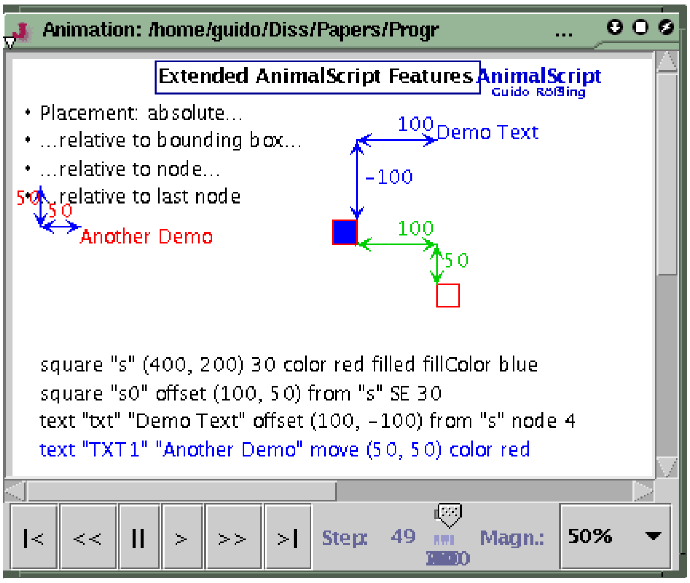


Fig. 2: Object placement using ANIMALSCRIPT: absolute coordinates, relative to bounding box or node, and as offset from the last node used

of where each object is at any given point in time by allowing direct access to its bounding box, the smallest rectangle containing the whole element. Users have the following choices for specifying the coordinates of any element:

- absolute coordinates, such as (400, 200),
- placement relative to a corner of a previously defined object. Any of the eight compass needle points plus the centre of the object's bounding box can be used as the reference point,
- placement relative to an arbitrary node of a previously defined polyline or polygon,
- placement relative to the last defined coordinate, which is helpful for operations such as "go forward 20 pixels",
- or relative to a previously defined location that may have been defined using any of the techniques listed above.

Figure 2 shows these applications and the ANIMALSCRIPT code required to produce them. Note that the filled square has been given absolute coordinates. The empty box's location is specified from an offset of the south east corner of the filled square and the text at the top right is placed relative to a node of the square. The last node used in the visualization was the top left corner of the text *"relative to last node"*, causing the text at the left centre of the screen to be placed relative to the text's top left corner.

As mentioned above, ANIMAL offers a method to translate all entries of the graphical user interface. ANIMALSCRIPT goes a step further by allowing the user to specify a translation for each text element. The languages supported in the visualization must be listed by language code at the start of the visualization file. Each text entry can then be given a translation prefixed by the language code. On loading the visualization, the user is prompted for the language to be used.

Note that the translation is highly unlikely to have the same length in pixels as the original element. For example, German texts tend to be longer than their English equivalent. Simply providing a translation is therefore not sufficient – elements that are supposed to be placed to the right of a translatable message would suddenly have an offset, or might overlap the message itself. By using relative placement aligned on the bounding box of the translated element, this problem can be solved.

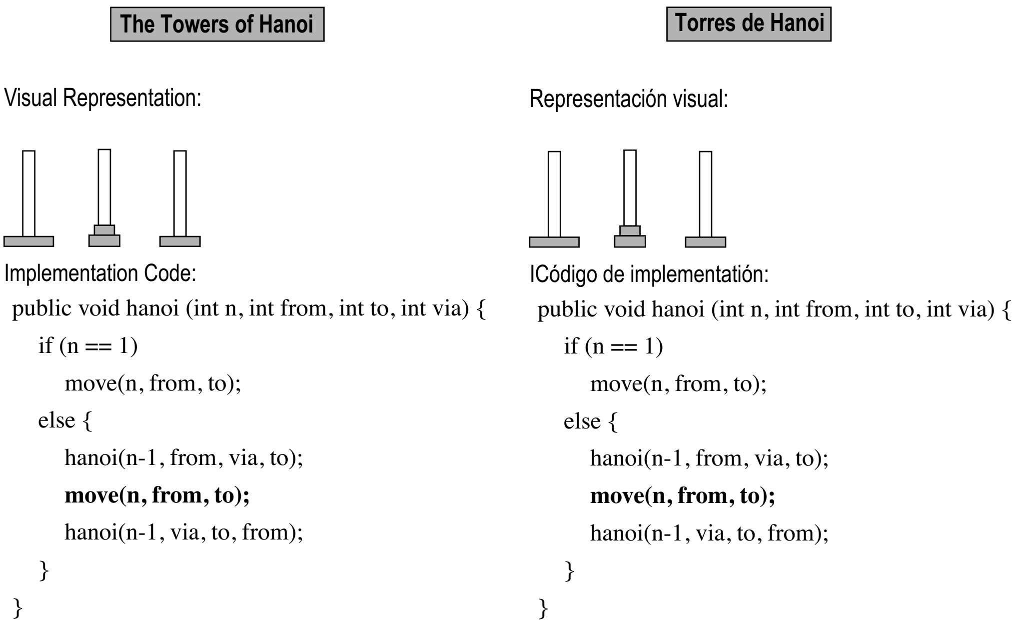


Fig. 3: Example visualization using the internationalization support of ANIMALSCRIPT

to Contents, page 1

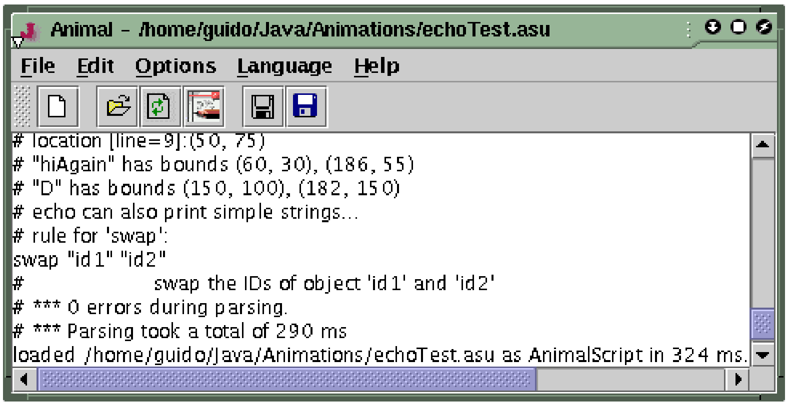


Fig. 4: Example applications of the echo command

After the text has been read in and translated to the chosen language, its location and extent is known, so that elements and effects depending on its placement can be correctly resolved. One implication of this is that users who work in an international context are not forced to generate separate visualizations according to the target audience. Instead, they can embed the translations into the same visualization and choose the language according to the context they are in at any given time. Figure 3 shows an example visualization with an English and a Spanish version. Note the changed width of the header rectangle that uses relative coordinates to fit under the heading.

The diagnostic output support of ANIMALSCRIPT is very helpful, especially for novice users. The functionality of ANIMALSCRIPT's echo command includes the printing of the location (upper left corner) or the bounding box of objects, the numbers of the components of a grouped object or currently visible objects, arbitrary text comments, or the rule syntax for ANIMALSCRIPT commands. Some of these applications are shown in Figure 4.

5 Extending ANIMALSCRIPT

We believe that ANIMALSCRIPT already contains several features that prove beneficial for users and allow the tool to be used in a generic context. However, we are of course aware that there are many application areas for which specific support would be helpful. For example, there is currently no support for graph structures or theoretical applications such as Turing machine or finite state automaton simulation. As we cannot possibly hope to satisfy all user wishes, we have adapted the structure of how ANIMALSCRIPT is parsed for easy extensibility.

Users who want to provide a specific extension to ANIMALSCRIPT should first check whether there is already a third-party implementation available. Otherwise, we can provide them with a set of comparatively simple steps to follow for providing extensions to ANIMALSCRIPT. To make this task as easy as possible, we provide a parser that translates a special subtype of Backus-Naur rules into Java code for parsing the commands and extracting properties. The user will then only have to provide some parts of the coding, particularly the mapping of properties into a set of ANIMAL transformations and primitive objects.

ANIMALSCRIPT parsing components are loaded dynamically from a configuration file. Therefore, users who have either implemented or downloaded an extension need only update the configuration file. No change to the ANIMAL system or ANIMALSCRIPT itself is necessary and no recompilation is required. This is also true for the process of extending the functionality of ANIMALSCRIPT, which does not require modifications or recompilation of existing code. The sole exception to this rule concerns changes to already implemented features, such as bug fixes.

6 Conclusions

In this article, we have outlined selected features of ANIMAL and ANIMALSCRIPT. ANIMAL is useful for generic software visualizations and does not require special user knowledge or skills due to the graphical editor. ANIMALSCRIPT adds powerful functionality for users willing to tackle animation generation by scripting, as is common in some other systems [Pierson/Rodger 98], [Stasko 98]. ANIMALSCRIPT currently supports only a limited number of operations. However, these operations already include advanced features that are not commonly found in many comparable applications, such as relative object placement, internationalization support and diagnostic output. In order to address as large a client audience as possible, both ANIMAL's GUI and ANIMALSCRIPT visualizations can easily be translated into other languages, especially when used in conjunction with relative object placement. Authors simply provide the translations without having to adapt object locations.

We hope that we will find users willing to help us add new features to ANIMALSCRIPT once we have documented how this can be achieved. Most changes do not require modifying ANIMALSCRIPT code. Further work in extending ANIMALSCRIPT includes support for images, specific computer science data structures or displays and the typical features found in presentation tools. In our experience, generating a visualization by scripting is much faster than doing the same visually, once the user has become familiar with the underlying concepts and commands. This is even true for conference presentations, despite the fact that most presentation tools rely heavily on graphical editing. Despite the lack of explicit ANIMALSCRIPT commands for item lists, we have used ANIMAL in conjunction with ANIMALSCRIPT as the main tool for conference presentations since February 2000 without problems or complaints.

References

[Ben-Ari 00]
M. Ben-Ari: An Extended Experiment with Jeliot 2000. First International Program Visualization Workshop, Porvoo, Finland, 2000.

[Haajanen et al. 97]
J. Haajanen, M. Pesonius, E. Sutinen, J. Tarhio, T. Teräsvirta, P. Vanninnen: Animation of User Algorithms on the Web. IEEE

Symposium on Visual Languages Proceedings (VL'97), Capri, Italy 1997, pp. 360–367.

[Hung/Rodger 97]
T. Hung, S. H. Rodger: Increasing Visualization and Interaction in the Automata Theory Course. 31st SIGCSE Technical Symposium on Computer Science Education, Austin, Texas 1997, pp. 6–10.

[Khuri/Hsu 00]
S. Khuri, H.-C. Hsu: Interactive Packages for Learning Image Compression Algorithms. 5th Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE 2000), Helsinki, Finland 2000, pp. 73–76.

[Naharro-Berrocal et al. 00]
F. Naharro-Berrocal, C. Pareja-Flores, and J. Á. Velázquez-Iturbide, Automatic generation of algorithm animations in a programming environment, 30th ASEE/IEEE Frontiers in Education Conference, Volume II, Stipes Publishing L. L. C., 2000, pp. 6–12 (session S2C).

[Pierson/Rodger 98]
W. C. Pierson, S. Rodger: Web-based Animation of Data Structures Using JAWAA. 29th SIGCSE Technical Symposium on Computer Science Education, Atlanta, Georgia 1998, pp. 267–271.

[Rößling et al. 00]
G. Rößling, M. Schüler, B. Freisleben: The ANIMAL Algorithm Animation System. 5th Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE 2000), Helsinki, Finland 2000, pp. 37–40.

[Silicon Graphics 98]
Silicon Graphics: JAL Algorithm Animation, 1998. Available online at http://reality.sgi.com/austern/java/demo/demo.html

[Stasko 98]
J. Stasko: Samba Algorithm Animation System, 1998. Available online at
http://www.cc.gatech.edu/gvu/softviz/algoanim/samba.html

[Zeller 00]
A. Zeller: Dynamic Display of Data Structures Using DDD. First International Program Visualization Workshop, Porvoo, Finland 2000.

to Contents, page 1

Automatic Web Publishing of Algorithm Animations

Fernando Naharro-Berrocal, Cristóbal Pareja-Flores, J. Ángel Velázquez-Iturbide, Margarita Martínez-Santamarta

We are undertaking a research action to simplifying the creation of program animations in order to foster a wider adoption of animations for CS education. In this article, we report on an extension of the WinHIPE functional programming environment to generate Web pages containing algorithm animations. Its integration within the programming environment guarantees the general applicability of the animator. In addition, by means of simple dialogs to specify the graphical style of visualizations, the animation interaction, etc., the generation of Web pages with animations is achieved automatically. We expect that the automatic nature of our extensions will encourage students to play an active role in learning and teachers to generate active documents containing active algorithm animations.

1 Introduction

In recent years, many animation systems have been developed for an educational use. Also on the Web, lots of pages containing collections of animations (e.g. [Brummund 97], [McCauley 00]) can be found. However, there are a number of practical problems in using animations extensively in education. Among these are:

- The difficulty of generating an animation. Programs must usually be modified, i.e. the user changes the program to generate the animation explicitly. Such modification is not simple: a new programming library must be learnt, and programs are complicated by annotations. While some systems include built-in animations, the user does not have the capability or the means to produce new ones.
- As empirical evidence shows, animations do not guarantee pedagogical success.

Empirical studies on the educational effectiveness of animations (summarized for example in [Ben-Ari 01], [Stasko/ Lawrence 98]) agree about several features:

- Close integration of algorithm animations and explanations. Otherwise, they are "nice movies".
- Rewind facility. When a student is uncertain about the behaviour of an algorithm, they should be able to backtrack and replay the animation.
- Active involvement of students. This can be achieved in several ways: allowing students to choose input data [Naps et al. 00], formulating "stop-and-think" questions [Naps et al. 00], fully integrating animation into the programming environment [Böcker et al. 86], and generating animations automatically [Naharro-Berrocal et al. 00].

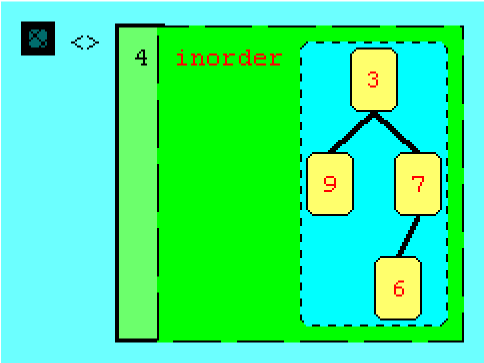
We have tried to meet these requirements by extending the capabilities of an integrated programming environment to the automatic generation of algorithm animations. In addition, the

Fernando Javier Naharro-Berrocal holds a M.Sc. degree from the *Universidad Politécnica de Madrid* (UPM). The *Ministerio de Educación y Cultura* granted him for study in the area of Architecture an Technology of Computer Systems. He was a *Profesor Asociado* in this area at the *Universidad de Castilla La Mancha* and since 1994 he is *Profesor Titular de Escuela Universitaria* in the area of Computer Languages and Systems at the UPM. He has been a researcher in several projects in these areas. His current interest consists in seeking innovation in the use of graphics in integrated development environments. <fnaharro@eui.upm.es>

Cristóbal Pareja-Flores holds a M.Sc. degree in Mathematics and a Ph.D. in Computer Science. Since 1998, he works as an associate professor at the *Dept. Sistemas Informáticos y Programación, Universidad Complutense de Madrid*. His interests include innovation in CS education and research on functional and concurrent programming, and programming environments. In these fields, he has collaborated and leaded several research projects and publications, national and international. He is the editor of the technical section "CS Education at the University" of Novática. <cpareja@sip.ucm.es>

J. Ángel Velázquez-Iturbide holds a M.Sc. degree (1985) and a Ph.D. in Computer Science from the *Universidad Politécnica de Madrid*. He is a *Profesor Titular de Universidad*, previously at the *Universidad Politécnica de Madrid* and since 1997 at the *Universidad Rey Juan Carlos*. He is the coordinator of the area of Computer Languages and Systems at his university. His interests are programming tools and environments, software visualization, and multimedia and Web applications, all of them mainly seeking innovation in computer science education. He is an associate editor of the "Computer Languages" section of Novática. He is a member of ACM, IEEE Computer Society, AACE and ATI. <a.velazquez@escet.urjc.es>

Margarita Martinez-Santamarta holds a M.Sc. degree and a Ph.D. in Computer Science from the *Universidad Politécnica de Madrid*. She is a *Profesor Titular de Universidad*, previously in the area of Computer Architecture and Technology of the *Universidad Politécnica de Madrid* and since 1998 in the area of Computer Languages and Systems of the *Universidad Rey Juan Carlos*. Her interests are human-computer interaction and parallel and distributed simulation. She has participated both as researcher and leader in several projects in these areas. She is a member of ACM, ATI, AIPO and UPA. <m.martinez@escet.urjc.es>



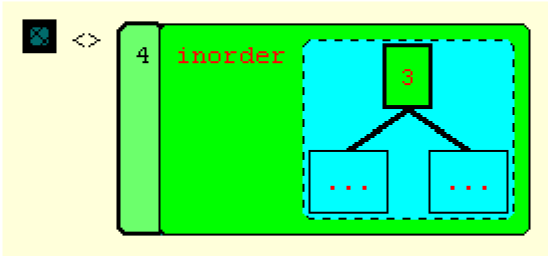


Fig. 1: Two different visualizations of an expression

integration of visualizations and animations into electronic documents should increase the understandability of algorithms. We report here on our efforts to automate the publishing of algorithm animations in Web pages. These can be used either by the teacher for lecturing or courseware, or by the student for experimentation and learning.

In the next section we give an overview of the programming environment *WinHIPE*. Section 3 describes the extensions to generate Web pages with animations. Finally, sections 4 and 5 contain related work and our conclusions.

2 *WinHIPE*: A Functional Programming Environment

WinHIPE is the *Windows* version of a programming environment for the functional language *Hope*. The design decisions underlying *WinHIPE* [Velázquez-Iturbide 94] have a pedagogical basis [Jiménez-Peris et al. 00]. In this section, we give an overview of its run-time model and its facilities to generate visualizations and animations.

2.1 Abstract Model of Expression Evaluation

Functional programs are executed by evaluating expressions. *WinHIPE* shows evaluation at a high level of abstraction: it models evaluation as a rewriting process [Velázquez-Iturbide 94], where the initial expression is rewritten into intermediate expressions until a final expression, i.e. the value, is obtained. An interesting feature of functional programming is that expressions contain both data and controls, so they are difficult to separate for visualization.

The programmer can control the pace of evaluation according to five options: to make 1 or *n* rewriting steps, or to evaluate the complete expression, a selected subexpression, or until a break-point is reached. An important design decision for the environment is that the whole expression is always shown, so the programmer always knows the evaluation state.

2.2 Visualization of Expressions

By means of simple dialogs, *WinHIPE* allows the programmer to customize the visualization of any intermediate expression in several ways [Velázquez-Iturbide/Presa-Vázquez 99]:

- The programmer can select either a pure textual or a mixed text-graphics visualization of expressions, where the graphical elements are lists or binary trees [Jiménez-Peris et al. 96].
- The typography of visualizations can be customized by tuning pretty-printing[1] formats, graphical elements, etc.
- Long expressions can be abbreviated automatically by a "focus+context" technique, only showing the parts most relevant to understanding of the current evaluation state.

Fig. 1 shows two different visualizations of an expression obtained while traversing a binary tree named "inorder". They exhibit different typographic elements, with the second expression being a simplification of the first.

2.3 Animation of Expression Evaluation

Animations are displayed in *WinHIPE* as a sequence of snapshots [Naharro-Berrocal et al. 00]. The raw material for building an animation is the visualizations of expressions generated by the programmer during an evaluation. At any moment, the user can build an animation by selecting a subset of the visualizations generated during the last evaluation. By default, the animation consists of all of the expressions involved, but the user may exclude irrelevant ones.

Note that the user can input any expression to evaluate and hence to animate (i.e. visualize in successive steps). Consequently, the system can animate any program. This contrasts with other systems where the user is restricted by either input data or the algorithm.

An animation is defined by means of a simple dialog asking for a directory to store the visualization and the location of a definition file. The definition file specifies whether the default is for the animation to be played automatically or to be controlled manually; in the former case, the speed of transition between snapshots must also be specified.

Depending on the definition file, when the user initiates an animation, it either starts running automatically or the first snapshot is displayed. In either case, the user uses an VCR-like interface: advance or backtrack one snapshot, play at the specified speed, pause, and go to the first or the last snapshot. The

1. A *pretty printer* is a program that formats a high-level program source text according to indentation rules based on the syntax and visualizing the program's structure.

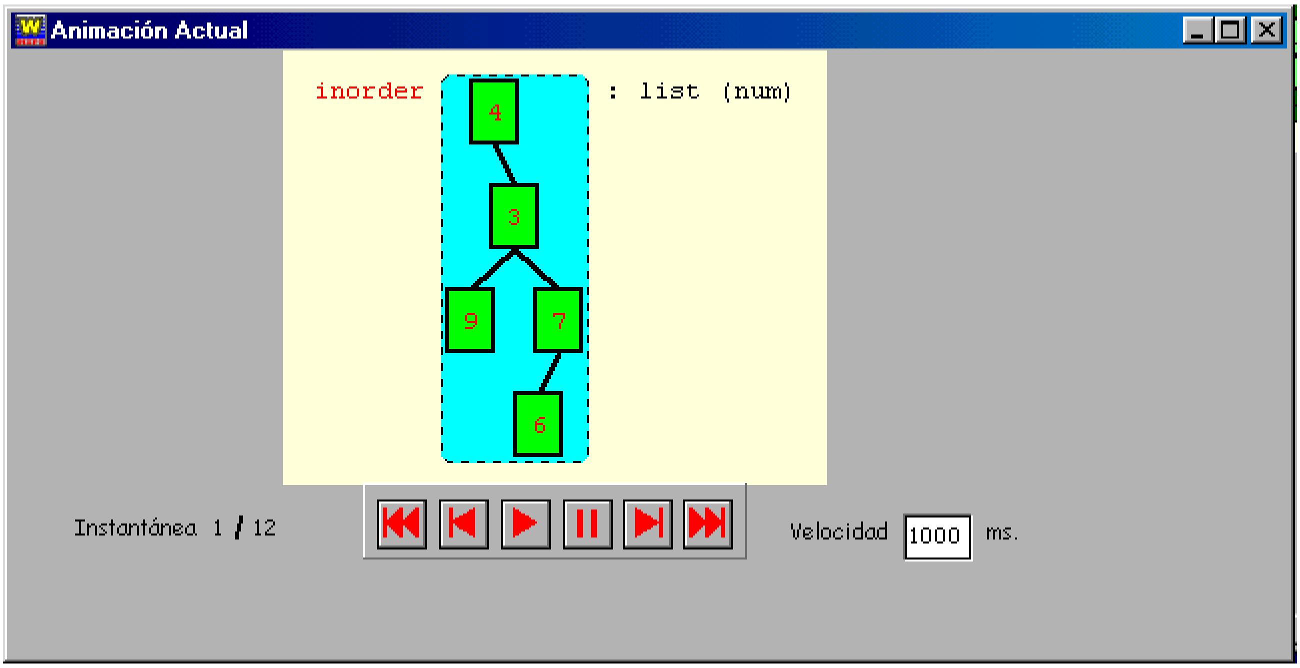


Fig. 2: interface of the animation player

number of the current snapshot relative to the total number of snapshots is also displayed (see Fig. 2).

3 Integration of Animations in Web Pages

We have extended *WinHIPE* to ennable automatic publication of Web pages with algorithm animations. In the following, we describe the Web pages and their generation process.

3.1 Structure of Animated Web Pages

Web pages are platform independent and readily integrate animations with explanations. Our Web pages contain the normal sections in algorithmic problem-solving, namely problem description, algorithm description, program text, and algorithm animation. Each section always has the same header (e.g. "problem statement"). The page also contains a local index for quick access to the different sections. The first two sections contain by default an inddication that no explanation is yet available; the user can edit these later manually. The Hope code used for the animation is included automatically in the third section from the programming environment. Finally, the animation is included with a control interface.

Fig. 3 shows a Web page for inorder traversal, where the animation is based on two simultaneous but differing snapshots.

3.2 Customization of Animated Web Pages

Animations within the programming environment show one snapshot at a time (see Fig. 2). However, other styles of interaction can also be useful, depending on the goal of the animation and the user's preferences. For instance:

- Two consecutive snapshots and an VCR interface. This is more helpful to a student to improve understanding of the transition from a state to the next.
- One snapshot and a sequence of indices, where each index corresponds to a different visualization. This is useful for a teacher who is fully conversant with an animation and wants direct access to snapshots during a lecture.
- All the snapshots in a single window controlled by a scroll bar.While this is not an animation *per se*, but the user can examine all the visualizations more flexibly.

We have provided the possibility of choosing the animation interaction by means of *style sheets*. A sheet style is a template of a dynamic HTML page that includes a JavaScript program to handle the animation interface. The template is parameterized with information specific to each animation: title of the animation, name of the Web page, number of visualizations in the animation, maximum size of the visualizations, directory where the visualizations are stored, and whether the animation will be shown in the same or in a new browser window.

A style sheet may also contain directives to include external files in order to allow more flexibility. Such directives are written within comments, analogous to compilation directives in some programming languages. For instance, the generated Web page might be based on a particular style sheet file. Or, the user might want to open a new browser window containing a given HTML file.

Style sheets and included directives are oriented to the automatic generation of Web pages. However, the user might want to modify manually the Web pages manually to improve them. Our technical choice of implementing style sheets with HTML and JavaScript permits this possibility because the code generated for the Web pages is open.

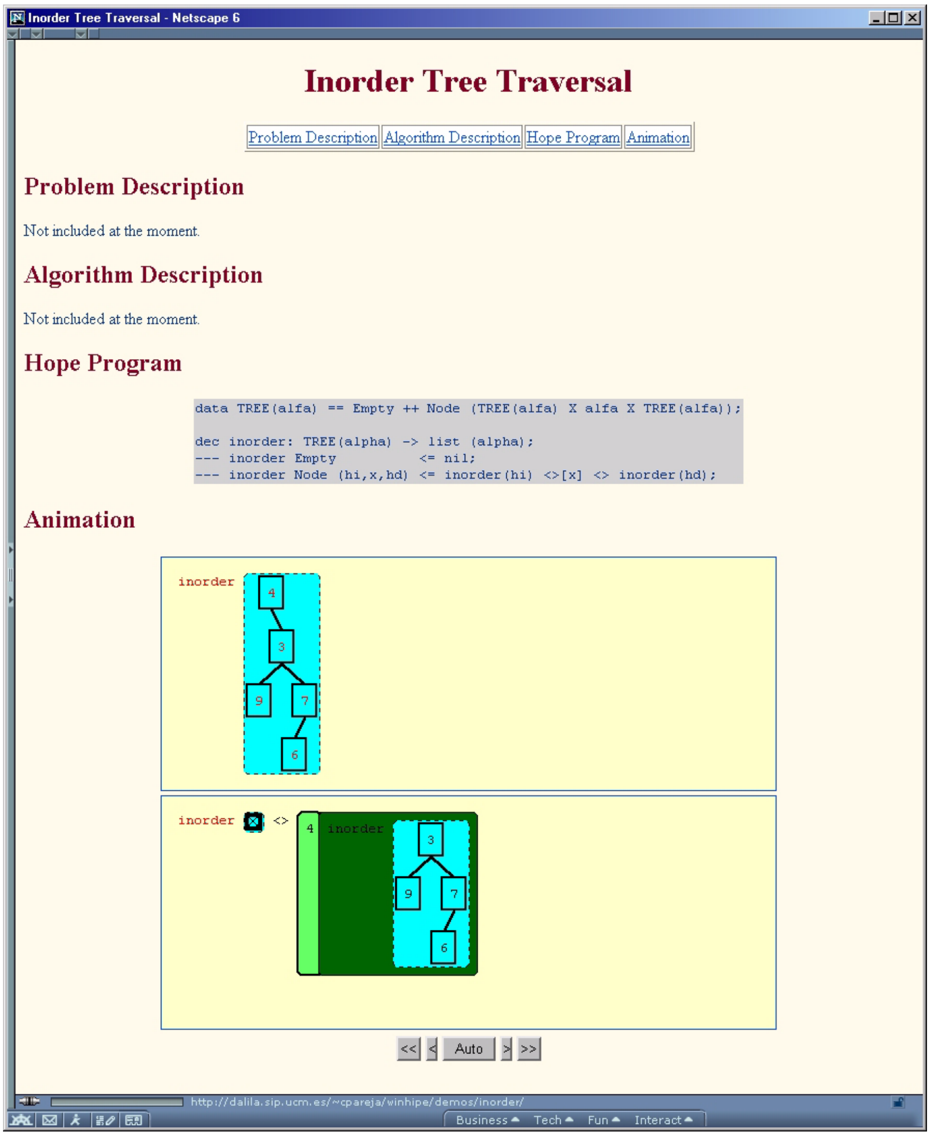


Fig. 3: A screen showing a generated Web page with an animation.

3.3 Generation of Web Pages

From a technical point of view, the animation is created as follows (Fig. 4 illustrates this process):

- The visualizations that will form the animation are selected, as explained in subsection 2.3.
- The options for generating a Web animation are selected, indicating the target directory, the style sheet, the title of the Web page, and the title of the explanation.
- The BMP files forming the visualizations in the original animation are converted into GIF files, according to the requirements of the style sheet and stored in the target directory. Depending on the requirements of the style sheet, it may also be necessary to extend the size of the smaller visualizations to match that of the largest one.
- Parameters in the selected style sheet are instantiated. Some parameters are computed automatically (number and size of visualizations, program code), but others depend on values supplied by the user in dialogs.
- External files in include directives are recursively instantiated and copied to the target directory.

4 Discussion and Related Work

We are aware of the fact that functional programming is a minority choice. Although it has only been adopted by a small number of universities, it is close to specification languages for abstract data types. From this point of view, animations generated by *WinHIPE* can be used in a data structures course. More formally, *WinHIPE* generates program (expression) visualizations, but their high-level of abstraction allows them to be considered as algorithm animations.

Recently, Tom Naps has proposed a framework to assess the pedagogical effectiveness of animation systems [Anderson/Naps 00]. He identified a set of features that constitute a design continuum for educational uses of animations. Since each feature cannot be marked with a single yes/no answer, our system *WinHIPE* meets (at least partially) all the features but the integration of animations and explanations. The current extension allows us satisf this requirement also.

An early unexpected result of our work was our interest in alternative interaction styles. When we identified several styles for Web pages, we realized that they are independent from the type of graphical window. As a result, they can also be useful for the programming environment itself.

In recent years, there has been considerable research on the development of animations for the Web. It is impossible to refer here to all of these efforts (read [Naps et al. 97] for a summary.) However, the need to build animations explicitly (and often Web pages)animations is a bottleneck for these systems.

5 Conclusions and Future Work

We have extended the functional programming environment *WinHIPE* to generate Web pages containing algorithm animations. The integration of this facility in the programming environment allows easy publication of animations for any algorithm. Visualizations, animations and Web pages are all generated automatically, only requiring the user to interact by means of simple dialogs. Of course, the user must plan careful-

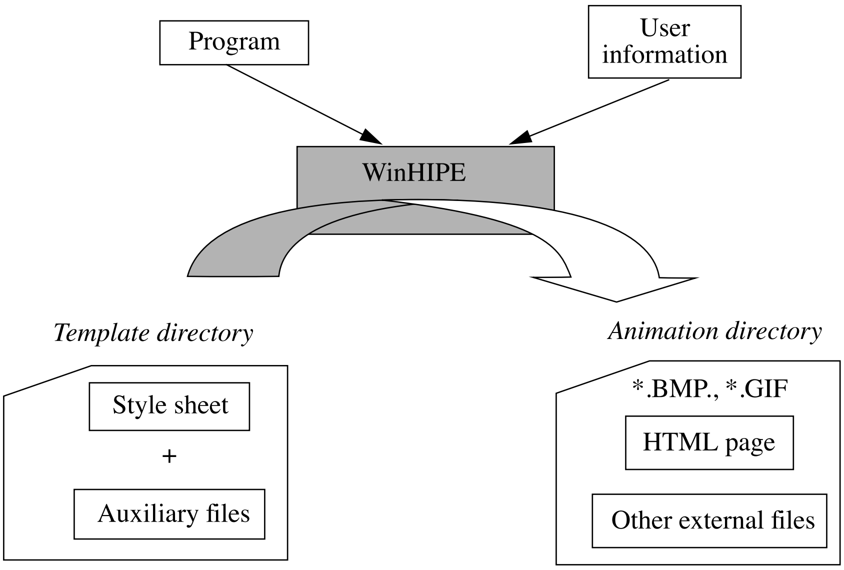


Fig. 4: Process of generating of Web pages with animations

ly the elements of an animation: typography of visualizations, input data for the algorithm, selection of snapshots, and interaction style. Style sheets have been developed with open code, so the user can easily modify them or create new ones.

From an educational point of view, our approach allows the integration of explanations and animations into Web pages. It provides a flexible interaction with the animation in a user friendly way (including a rewind operation). The automatic nature of our extensions encourages the student to play an active role in learning and the teacher to generate active documents containing algorithm animations.

We are developing a Web site with a collection of animations generated with *WinHIPE* (currently in Spanish) at http://dalila.sip.ucm.es/~cpareja/winhipe/demos/demos.html.

We are improving the animation features of our programming environment and attempting to export them to the Web as much as possible. Initially, we would like to produce explanations associated to individual visualizations automatically to complement the static one provided currently by the user. Second, we are improving facilities for navigation and customization of animations. Finally, we plan to measure the learning effectiveness of alternative interaction styles in order to use them effectively in education.

Acknowledgments

This work was supported by projects 07T/0036/98 of the Comunidad Autónoma de Madrid and TIC2000–1413 of the CICYT agency.

References

[Anderson/Naps 00]
J. M. Anderson and T. L. A. Naps: A context for the assessment of algorithm animation systems as pedagogical tools, ACM Program Visualization Workshop (Porvoo, Finland, 2000), position statement.

[Ben-Ari 01]
M. Ben-Ari: Program visualization in theory and practice, Informatik/Informatique, this issue.

[Böcker et al. 86]
H.-D. Böcker, G. Fisher, and H. Nieper: The enhancement of understanding through visual representations, ACM CHI'86 Conference, 44–50.

[Brummund 97]
P. Brummund: The Complete Collection of Algorithm Animations, http://www.cs.hope.edu/~alganim/ccaa/ccaa.html, 30th July 1997.

[Jiménez-Peris et al. 96]
R, Jiménez-Peris, C. Pareja-Flores, M. Patiño-Martínez and J. Á. Velázquez-Iturbide: Graphical visualization of the evaluation of functional programs. Proc. ITiCSE'96 (Barcelona, Spain, 1996), 36–38.

[Jiménez-Peris et al. 00]
R. Jiménez-Peris, C. Pareja-Flores M. Patiño-Martínez and J. Á. Velázquez-Iturbide: Towards truly educational programming environments. T. Greening (ed.), Computer Science Education in the 21st Century, Springer-Verlag, New York, 2000, 81–112.

[McCauley 00]
McCauley, R. Computer Science Education Links, section Algorithms and Algorithm Animation, http://www.cacs.usl.edu/~mccauley/edlinks/#Algorithms, 12th February 2000

[Naharro-Berrocal et al. 00]
F. Naharro-Berrocal C. Pareja-Flores and J. Á. Velázquez-Iturbide: Automatic generation of algorithm animations in a programming environment. Proc. 30th ASEE/IEEE Frontiers in Education Conference (Kansas City, Missouri, USA, 2000), 6–12 (session S2C).

[Naps et al. 97]
T. L. Naps et al.: Web based visualization. Proc. ITiCSE'97 (Uppsala, Sweden, 1997).

[Naps et al. 00]
T. L. Naps J. R. Eagan and L. L. Norton: JHAVÉ: An environment to actively engage students in Web-based algorithm visualizations. Proc. 31st SIGCSE Tech. Symp. (Austin, Texas, USA, 2000) 109–113.

[Stasko/Lawrence 98]
J. T. Stasko and A. Lawrence: Empirically assessing algorithm animations as learning aids. Stasko, J.T. et al. (eds.), Software Visualization, MIT Press, 1998, 419–438.

[Velázquez-Iturbide 94]
J. Á. Velázquez-Iturbide: Improving functional programming environments for education. M.D. Brouwer-Janse and T.L. Harrington (eds.), Man-Machine Communication for Educational Systems Design, Springer-Verlag, Berlin, 1994, 325–332.

[Velázquez-Iturbide/Presa-Vázquez 99]
J. Á. Velázquez-Iturbide and A. Presa-Vázquez: Customization of visualizations in a functional programming environment. Proc. 29th ASEE/IEEE Frontiers in Education Conference (San Juan, Puerto Rico, 1999), 22–28 (session 12b3).

Parameters Affecting QoS in Voice over Packet Networks

Antonio Estepa, Rafael Estepa and Juan M. Vozmediano

This article presents an overview of the factors that negatively affect speech quality in the increasingly popular area of voice over IP networks. These factors are packet network-related as well as terminal-related. Methods available to control speech quality within each of these areas are considered.

Keywords: QoS, terminals, VoIP

1 Introduction

Traditionally, voice and data networks have been supported by two separated infrastructures. The voice transport over IP (VoIP) offers service providers the advantage of integrating their networks into a single infrastructure. The substitution of traditional voice networks by VoIP will depend on the capability to offer to the end user a similar quality.

The speech communication quality is a subjective concept: it depends on the level to which users' expectations are met. Opinions on the quality of a call are based on several factors related to the subjective perception of the call: overall quality, volume, intelligibility, speaker identification and naturalness. Physical factors such as loss, echo, delay and noise also affect the perception of the speech [Rix et al. 00].

These factors are physically located in three points: terminals, the IP network and when needed, in interconnection devices between the IP and PTSN networks called gateways. This article addresses the call quality in a scenario with PC-based terminals where the voice packets traverse only an IP network.

The paper is organised as follows. The next section describes the main factors that impact the call quality. Later, quality aspects are analysed within the IP network and terminals in sections 3 and 4 respectively. Finally, section 5 concludes the paper.

2 Impairments Affecting the Call Quality

There are several well-known impairments that have traditionally been associated with speech quality in plain old telephone services (POTS). The main impairments come from [ETSI 96]:

- *Loudness Ratings*. Loudness ratings are objective measures of the loudness loss, i.e. a weighted, electro-acoustic loss between certain interfaces in the network. Usually expressed in dB, the Overall Loudness Rating (OLR) defines the sound power loss between the acoustical signal at the microphone and the listener's ear.
- *Echo*. The echo is a copy of the original sound with a delay long enough to be perceived as a second, undesired, sound. The negative impact of the echo will depend on both the power and delay of this second sound. The talker can receive the echo of his own voice (talker echo), while the listener can receive multiple copies of the original sound differently delayed (listener echo). The main echo sources in VoIP are focused in terminals and will be examined in the terminals QoS section.
- *Sidetone*. The sidetone is the sound emanating from the surroundings and perceived by the listening ear at the same time as the speech. It can be decomposed into ambient noise (listener sidetone) and the sound of the listener's voice (talker sidetone). Note that only the listening ear is considered, as the brain hearing mechanism ignores the free ear. Talker and listener sidetones are characterized in the form of Loudness Ratings.
- *Delay*. The delay is usually higher in packet switching networks than in circuit networks, so it becomes an important impairment in the VoIP QoS. The packet network, terminals, and gateways will add delay to the communication. The additive effects of these delays will impose a limit on the achievable quality [Vleeschauwer et al. 00]. According to [ITU 96], delays longer than 400 ms will cause a lack of interactivity in conversations.

  Delays take place in:
  - *IP Networks*. Network delay in packet networks is due to transmission delay in every link along the path to the destination and queuing delay in every router. Transmission delay depends on subnetwork technology, while queuing delay depends on queuing policies at the routers.

Juan M. Vozmediano received the M.S. and Ph.D. degrees in telecommunications engineering from the Technical University of Madrid, Spain, in 1989 and 1994 respectively. In 1993, he joined the Area of Telematic Engineering of the University of Seville, as Associated Professor. <jvt@trajano.us.es>

Antonio Estepa-Alonso received the M.S. degree in telecommunications engineering from the University of Seville, Spain, in 1999. Since 2000, he has been working as Assistant Professor in the Area of Telematic Engineering of the University of Seville. <aestepa@trajano.us.es>

Rafael Estepa-Alonso received the M.S. degree in telecommunications engineering from the University of Seville, Spain, in 1998. Since 1999, he has been working as Assistant Professor in the Area of Telematic Engineering of the University of Seville. <rafa@trajano.us.es>

- *Terminals*. Terminal delay will strongly depend on terminal type and processing power. Software-based terminals will show higher delay than hardware-based ones.
- *Special Equipment and Noise*. Special equipment such as low rate codecs is necessary to reduce network load, but this will also degrade speech quality by increasing quantizing noise, distortion and delay.
  The noise sources come from the power sum of electric circuit noise, and room noise. Noise sources at each end of the communication can be easily controlled by proper tuning of the loudness ratings and listener sidetone.

The network basically affects the network delay (and jitter) and also the packet loss, but, as delay and packet loss in the network are correlated [Jiang/Schulzrine 99], it could be stated that a network with a limited delay will also have a negligible packet loss. In the next section, the methods available to control packet loss and delay are examined.

Terminals contribute to increase the delay: they are also the main source of echo in the system. Sidetone, low rate codecs, and quantization distortion also impair the quality perceived by the users.

3 QoS in the IP network

IP offers an unreliable, connectionless network-layer service that is subject to packet loss, reordering and duplication, all of which, together with queuing delay in router buffers, will increase with network load. Because of the lack of any firm guarantees, the traditional IP delivery model is often refered to as "best-effort" and an additional higher layer end-to-end protocols such as TCP is required to provide end-to-end reliability. As the traffic in the network increases, network service degrades gracefully causing problems for applications with real-time requirements such as telephony.

QoS protocols provide additional capabilities to the network allowing it to distinguish traffic with real-time requirements from that which can tolerate delay, jitter and loss.

There are two strategies for QoS provisioning:

- Resource Reservation (*Integrated Services*): network resources are apportioned according to an application's QoS request and subject to a bandwidth management policy.
- Prioritization (*Differentiated Services*): network traffic is classified and apportioned network resources according to bandwidth management policy criteria. To enable QoS, network routers give preferential treatment to classifications identified as having more demanding requirements.

These strategies can be applied to individual application flows or to flow aggregates. A flow is defined as an individual, uni-directional data stream between two applications (sender and receiver) uniquely identified by a 5-tuple (transport protocol, source address and port number and destination address and port number). Two or more flows with something in common are named aggregates.

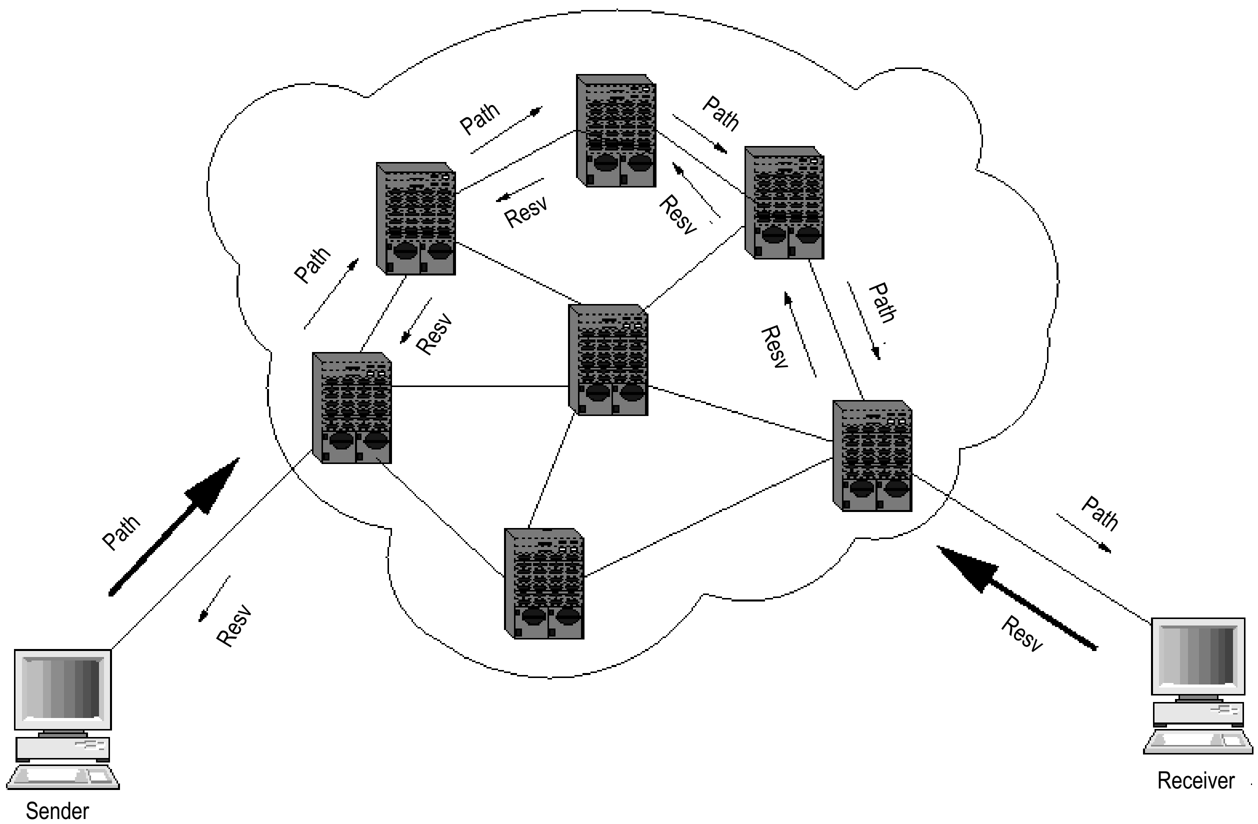


Fig. 1: RSVP End-to-End

Applications, network topology and policy dictate which type of QoS is most appropriate in a given situation. To accommodate the need for these types of qualities there are different protocols and algorithms:

- Reservation Protocol (RSVP): provides signalling to enable Integrated Services. Typically used in a per-flow basis.
- Differentiated services (DiffServ): provides a coarse and simple way to categorize and manage network traffic (flow) aggregates.
- Multi Protocol Label Switching (MPLS): provides bandwidth management for aggregates via network routing control according to labels in (encapsulating) packet headers.

These protocols are introduced in more detail in the next subsections.

3.1 Integrated Services (IntServ) and RSVP

Integrated Services enable the coexistence of best-effort datagram delivery and enhanced quality of service delivery classes with regard to bandwidth, packet queuing delay and loss. The level of QoS provided by these enhanced classes is programmable on a per-flow basis according to requests from the end applications. These requests can be passed to the routers by network management or, more commonly, using a reservation protocol such as RSVP.

IntServ basically offers two additional delivery services (as well as the traditional best-effort) to end applications:

- Guaranteed Service: this service guarantees bandwidth as well as establishing a bound on end-to-end delay which is very important for the VoIP applications.
- Controlled Load: this is equivalent to "best effort service under unloaded conditions". It is better than best-effort but cannot provide the strictly bounded service that guarantees the real-time requirements.

For traffic control, IntServ performs: admission control, packet classifying, packet scheduling and a setup protocol (RSVP) [IETF 94].

RSPV is a signalling protocol that provides reservation setup and control and enables the integrated services (IntServ) [IETF 97].

RSVP-enabled senders characterize outgoing traffic in terms of the upper and lower bounds of bandwidth and jitter. RSVP senders periodically transmit PATH messages that contain this traffic specification (Tspec) information to the (unicast or multicast receiver(s)) destination address. Each of any RSVP-enabled router along the downstream route establishes a "path-state" that includes information from the previous source address of the PATH message.

To make a resource reservation, receivers send a RESV (reservation request) message upstream. In addition to Tspec, the RESV message indicates the type of integrated service required either Controlled Load or Guaranteed, indicating a filter specification (filter spec). VoIP packets could then indicate the bounds for the delay and jitter of the communication.

When an RSVP router along the upstream path receives the RESV message, it uses the admission control process to authenticate the request and allocate the necessary resources for that flow and sends the RESV upstream to the next router. Reservations in each router are "soft", which implies the need for periodical refreshing. For scalability reasons, this mechanism is not recommended for the core routers of the network.

When the last router receives the RESV and accepts the request, it sends a confirmation back to the receiver.

Although RSVP traffic can traverse non-RSVP routers, this creates a weak-link in the chain where service falls-back to best-effort. Another weakness of this protocol is that reservations are receiver-based and it does not consider special mechanisms for multicast receiver groups.

RSVP allows an application to request QoS with a high level of granularity and with the best guarantees of service delivery. All these possibilities come at the price of complexity and overhead.

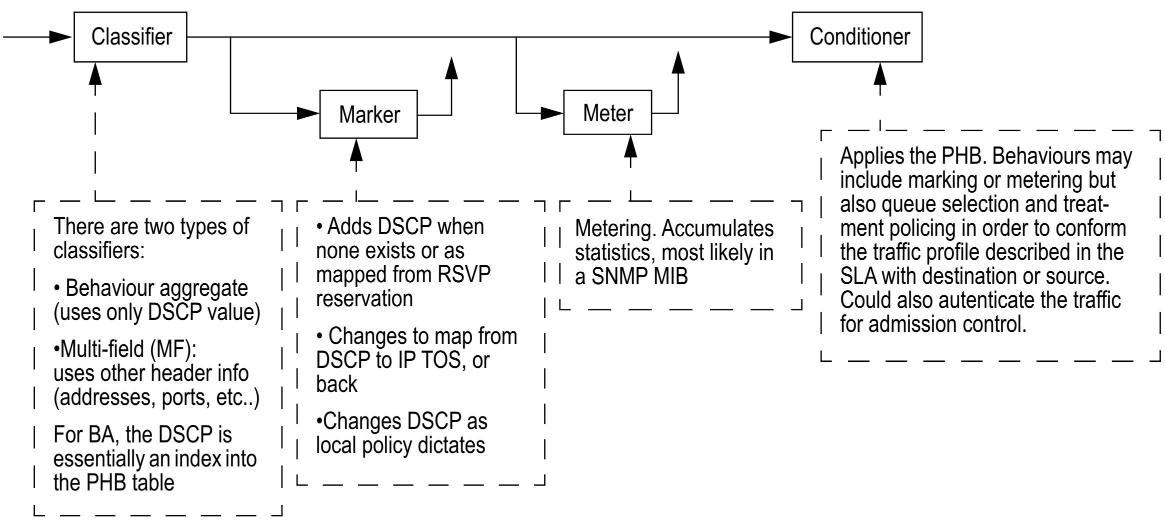


Fig. 2: Differentiated Services architecture. This functionality is enabled in every DiffServ enabled router, although not all functions are used all the time.

3.2 Differentiated Services (DiffServ)

Differentiated Services provides a simple method of classifying services offered to various applications. Service classes are identified, packets are marked as belonging to a particular service sent on its way and routers in the path examine headers to determine the treatment for the aggregate flow [IETF 98].

There are currently two standard per hop behaviours (PHBs) defined that effectively represent two service levels:

- Expedited Forwarding (EF): it has a single codepoint (Diffserv value). EF minimizes delay, jitter and provides the highest level of aggregate quality of service. Any traffic that exceeds the traffic profile (which is defined by local policy) is discarded.
- Assured Forwarding (AF): it has four classes and three drop-precedences within each class (a total of twelve codepoints). Excess AF traffic is not delivered with as high probability as indicated in the traffic profile, which means that it may be delayed, but not necessarily dropped.

As illustrated in figure 2, PHBs are applied to traffic by the conditioner at a network ingress point (network border entry) according to pre-determined policy criteria. The traffic may be marked at that point, and routed according to the mark, then unmarked at the network egress point.

Diffserv assumes the existence of a service level agreement (SLA) between adjacent networks. The SLA establishes the policy criteria, and defines the traffic profile. It is expected that it will be policed and smoothed at egress points according to the SLA, and any traffic above the upper bound at an ingress point, is not guaranteed (or may incur extra costs, according to the SLA).

When applied, the protocol mechanism makes use of bits in the "DS-byte". The DS-byte denotes the service that the packet should receive, and travels in the Type Of Service field in IPv4 and in the Traffic Class Octet field in IPv6.

The simplicity of DiffServ to priorize traffic belies its flexibility and power. In contrast with RSVP, the amount of state information depends on the number of classes, not the number of flows. Sophisticated classification, authentication and, marking and shaping operations are only needed at boundaries and it is the sender who requests resources, not the receiver.

3.3 MPLS

Multi-Protocol Label Switching (MPLS) works by building engineered paths across the core of an IP network, then sending packets along those predefined paths.

When a packet enters the network, an edge route looks up the destination address of the packet and tags it with a label that specifies the route and, optionally, class of service (CoS) attributes. The idea of MPLS is that by using the label to determine the next hop, routers have less work to do and can act like simple switches. As the labelled packet moves across the network, each router uses the label to choose the destination, and optionally the CoS, of the packet, rather than looking up the destination address for each packet in a routing table. As the packet leaves the core of the network, an edge router uses the destination address to direct the packet to its final destination. Subsequent packets in the data stream are quickly and automatically labelled in this way, as they can be anticipated.

Label switch routers (LSRs) build the path that a packet takes across the core of the network, called a label switched path (LSP). Labels stored by each router define the path, which can follow specific routes or constraints. LSRs at the core of the network participate in routing topology exchanges and become true peers with the edge routers. The number of peers each edge router must communicate with is reduced to the immediately adjacent LSRs and routers.

Labels can be used to identify traffic that should receive special treatment to meet QoS requirements. By using sophisticated traffic management techniques for the LSPs defined by the LSR, guaranteed service level agreements can be delivered in an IP network environment.

A more complex aspect of MPLS involves the distribution and management of labels among MPLS routers, to ensure they agree on the meaning of various labels [Rix et al. 00]. The Label Distribution Protocol (LDP) is specifically designed for this purpose, but it is not the only possibility.

3.4 QoS Architectures

In real-world use, it is unlikely that these protocols will be used independently, and in fact they are designed for use with other QoS technologies to provide end-to-end QoS between senders and receivers.

IntServ/DiffServ: RSVP provisions resources for network traffic, whereas Diffserv simply marks and prioritizes. RSVP is more complex and demanding than DiffServ in terms of router requirements, so it can negatively affect the backbone routers. This is why DiffServ is preferred in the backbone.

DiffServ is a perfect complement to RSVP, as the combination enables end-to-end QoS. End hosts may use RSVP request with high granularity (e.g. bandwidth, jitter threshold, delay …). Border routers at backbone ingress points can then map those RSVP reservations to a class of traffic indicated by a DS-

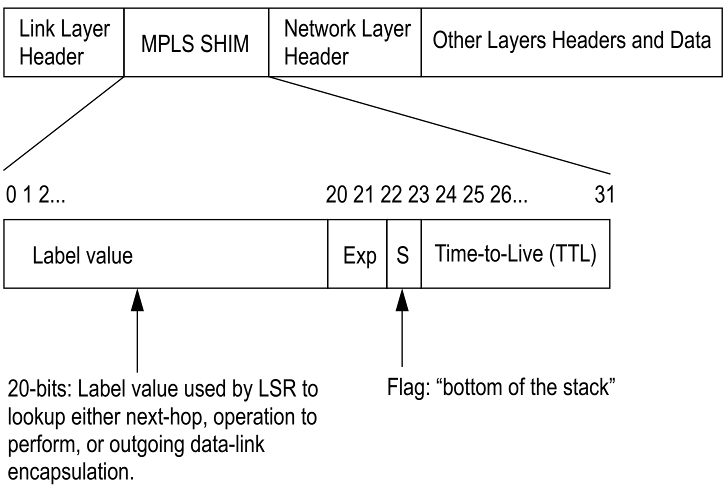


Fig. 3: MPLS label stack entry used to "encapsulate" IP header.

byte. At the backbone egress point, the RSVP provisioning may be honoured again, to the final destination.

MPLS and IntServ/DiffServ: in coexistence with RSVP, it is possible for MPLS to assign labels according to the RSVP flows.

MPLS and DiffServ are similar with respect to the qualitative QoS support, so the mapping of DiffServ traffic to MPLS tags is relatively simple. Supporting the DiffServ per-hop model also requires the allocation of a set of aggregate forwarding resources for each DiffServ forwarding class in each LSR. Additionally, an LSR may need to associate the packet with a particular drop precedence.

4 QoS in terminals

As seen above, QoS mechanisms allow network delay and jitter to be controlled. However, terminals collect the rest of the impairments of the calls.

Terminals can be classified into digital terminals and adapters for traditional analog terminals. Digital terminals can be implemented as software programs running on a PC (PC-based terminals) or as hardware devices (cards, IP telephone sets). According to the audio system used, digital terminals can be further classified into handset terminals, headset terminals and speaker phone terminals.

4.1 Echo

In a scenario with two PC-based terminals, the main echo source will be the acoustical path, as all the circuits are 4-wire. Echo sources will always be located at terminals unless there are gateways to circuit switching networks.

A PC terminal with loudspeakers and a microphone will be severely affected by acoustic echo, with a negligible echo loss. Stability problems may arise, as the speakers can amplify the signal up to the saturation point. For digital terminals with handsets or headsets, typical echo loss values are greater than 45 dB [Janssen et al. 00], which assures a negligible echo. For speaker phone terminals, echo loss decreases to 15 dB (even lower than in PTSN) so exhaustive acoustic echo control is necessary in the terminal.

4.2 Delay in Terminals

Whichever the type, terminal delay will be due to the following elements:

- Codecs: they will add an algorithmic delay, depending on the processing power and codec type; a prediction delay, due to look-ahead algorithms; and a frame-filling delay due to packeting time. For instance, look-ahead delay for G.723.1 codecs is 7.5 ms, and frame delay in G.723 is 30 ms.
- Protocol overhead: each protocol layer will add headers and processing time. Delay due to protocol overhead will depend on the number of codec frames per IP packet. With few frames per IP packet, the overhead will be high. With many frames, the protocol overhead will be smaller but the filling-up time will increase.
- Transmission delay: at the terminal, this delay will linearly increase with packet size.
- Buffering delays due to devices associated with the operating system, such as buffers to handle the sound card, network interface and the used client application.
- Playout Buffer: upon receiving a frame, it will be placed into the jitter buffer. The frame will be delayed an amount of time depending on the jitter control needs. Packets arriving later than their scheduled playout time (jitter buffer delay included) will be discarded and it will result into an increment of packet loss. However a big buffer size will derive into a growth of the overall delay.

4.3 Special Equipments: Low Rate Codecs

Almost any type of low rate codec [Hersent et al. 00] will increase the distortion and quantizing noise when compared to G.711 coding. Their response can also be affected by voice signal level, voice tone, room noise, transmission errors and delay variations.

Moreover, the impact of packet loss will depend directly on the codec. While a single lost packet may be unnoticed in G.711 coding, it may affect a rather long piece of speech in low rate codecs. Effects of low rate codecs are usually perceived as voice clipping [ITU 99].

4.4 Loudness Loss and Sidetone

In a typical sound card the sidetone path goes from the microphone to the loudspeaker. As this path is via hardware, the delay of the sidetone signal is not longer than 2 ms, so it can not be confused with an echo signal.

The microphone gain of the sound card can be set to optimize the talker sidetone loss (STMR). Sidetone in PC-based terminals should be tuned to approximate STMR values of 15dB.

With regard to loudness loss in VoIP terminals, the gains can be easily tuned by the users themselves (i.e. values tuned to a level comfortable for users). Thus, theoretically, the optimal values for the end-to-end communication can be chosen. Optimal values will also minimize other effects such as the echo.

5 Conclusions

Traditional impairments sources in POTS have evolved in the VoIP scenario. In VoIP calls, delay (and jitter) and acoustic echo are the main factors determining the quality. Special devices such as low rate codecs should also be considered, specially when packet losses are occurring.

In this article, impairments and methods to attenuate their influence have been examined. In section 3, quality aspects referred to the network have been addressed. Mixed architectures can provide the IP-network with extra capabilities that meet the requirements of real-time applications.

In the last section, QoS factors within terminals have been described. Delay and acoustic echo have been examined in more detail. The terminal delay can be kept under control, but echo must be controlled by local echo control procedures when necessary.

The existence of QoS mechanisms to limit delay, packet loss and jitter in the network; and the possibility of controlling the echo at the terminals, can facilitate a fast deployment of VoIP.

to Contents, page 1

Acknowledgements

The work leading to this article has been partly supported by CICYT and the EU under contract number 1FD97-1003-C03-03.

References

[ETSI 96]
ETSI ETR 250. Speech communication quality from mouth to ear for 3,1 kHz handset telephony across networks. (7/96).

[Hersent et al. 00]
O. Hersent, D. Gurle, J.P. Petit. IP Telephony. Addison-Wesley. 2000. ISBN 0201-61910-5.

[IEC]
Multiprotocol label Switching (MPLS). The International Engineering Consortium.

[IETF 94]
IETF RFC 1633. Integrated Services in the Internet Architecture: an Overview. (6-1994)

[IETF 97]
IETF RFC 2205. Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification. (7-1997)

[IETF 98]
IETF RFC 2475. An Architecture for Differentiated Services. (12-1998)

[ITU 99]
ITU-T Recommendation G.113 Appendix I. Provisional Planning Values for the Equipment Impairment Factor Ie. (09/99)

[ITU 96]
ITU-T Recommendation G.114. One-Way Transmission Time. (02/96)

[Janssen et al. 00]
J. Janssen, D. Vleeschauwer, G.Petit. Delay and Distortion Bounds for Packetized Voice Calls of Traditional PSTN Quality. IPTEL 2000, (12–13/4/00) Berlin.

[Jiang/Schulzrine 99]
W. Jiang, H. Schulzrine. QoS Measurements of Internet Real-Time multimedia Services. (12/99)

[Rix et al. 00]
A. Rix, J. Beerends, M. Hollier, A. Hekstra. PESQ – the new ITU standard for end-to-end speech quality assessment. AES 109th CONVENTION. LOS ANGELES, 2000 SEPTEMBER 22–25.

[Vleeschauwer et al. 00]
D. Vleeschauwer, J.Janssen, G.H. Petit, F. Poppe. Limites de calidad para el transporte de paquetes de voz. Alcatel Telecommunications Review. First quarter 2000.

to Contents, page 1

Signalling in Voice over IP Networks

José Ignacio Moreno, Ignacio Soto and David Larrabeiti

Voice signalling protocols have evolved, keeping with the prevalent move from circuit to packet switched networks. Standardization bodies have provided solutions for carrying voice traffic over packet networks while the main manufacturers are already providing products in workgroup, enterprise, or operator portfolio. This trend will accrue in next years due to the evolution of UMTS mobile networks to an "all-IP" environment. In this paper we present the various architectures that are proposed for signalling in VoIP, mainly: H.323, SIP and MGCP. We also include a brief summary about signalling in classical telephone networks and, at the end, we give some ideas about the proposed "all-IP" architectures in UMTS 3G mobile networks.

1 Signalling in telephone networks

Signalling in classic telephone networks has evolved dramatically during the 20th century, at the pace driven by the development of circuit switching technologies these networks are based on. The manual switches in use at the end of the 19th century were replaced by electro-mechanical switching in the advent of the 20th century. This technological stage would last until the 60s. Signalling was carried "in-band" (level change and tones inside the telephone channel), and was interpreted by electromechanical and electronic elements (relays and filters) on its way through the network. End-to-end physical connectivity quickly evolved to logical connectivity, and transmission, formerly analog FDM, moved to digital TDM structured as 64 Kb/s channels.

In the middle 60s, transmission and switching integrated digital network merged with the advent of digital exchanges and CPU-controlled switching (Stored Program Control concept). The 64 Kb/s synchronous channels are byte-by-byte switched in space and in time. These exchanges are integrally controlled by processors that exchanges a signalling protocol with other exchanges' processors. The first signalling protocols used in these systems were based on the status of a few bits in the TDM frame permanently attached to each voice channel, just as binary representations of predecessor analog systems. The quantum leap in the history of telephone signalling was the application of computer networks technology to the design of the signalling system. Signals became messages exchanged by systems over an independent packet switching network exclusively dedicated to this task.

Although this type of operation is now almost ubiquitous in the public telephone network, the last segment to be digitised – the subscriber loop – remains analog, with a small deployment of fully digital accesses (ISDN). As a result, existing user-network signalling has evolved very little as compared to the revolution observed in the network-to-network architecture, that has enabled the development of many additional services, mobile telephony, intelligent network services, broadband ISDN and internetworking with VoIP systems.

The network-to-network signalling system that has supported this evolution is the *Signalling System number 7* (SS7). The first CCITT standard is dated 1981 ("Yellow Book"), and has been refined and enhanced in successive editions in 1985 ("Red Book"), in 1989 ("Blue Book") [SS7 89] and following ITU-T standards.

SS7 is a complete protocol stack where signalling units are messages issued by signalling applications transported in packets. The essential features of this system are:

- Bundles of signalling links and nodes build a packet switching network that is logically independent from the circuit switching one, with a specific numbering plan and internationally defined by ITU-T.
- SS7 is a common channel signalling system. A set of channels between Signalling Points at exchanges (and Signalling Transfer Points) is dedicated to transport signalling to setup, release, supervise, etc. any voice or data 64 Kb/s channel. In previous signalling systems, a signalling associated to each voice circuit was transported over a transmission channel exclusively dedicated to it.
- SS7 is a 4-level protocol stack (Figure 1).

Jose Ignacio Moreno is Assistant Professor of Computer Networking at *Universidad Carlos III* of Madrid. He obtained his Ph.D. at the Technical University of Madrid in 1996. From 1991, he has participated in several national and international research projects in the field of telematic services and networks. <jmoreno@it.uc3m.es>

Ignacio Soto is Associate Professor of Computer Networking at *Universidad Carlos III* of Madrid. He obtained his Ph.D. at Vigo University and worked in different National and European research projects. <isoto@it.uc3m.es>

David Larrabeiti is Assistant Professor of Computer Networking at *Universidad Carlos III* of Madrid. He obtained his Ph.D. at the Technical University of Madrid in 1996. From 1991, he has participated in several European research projects related to the study and deployment of next generation networks and services. <dlarra@it.uc3m.es>

to Contents, page 1

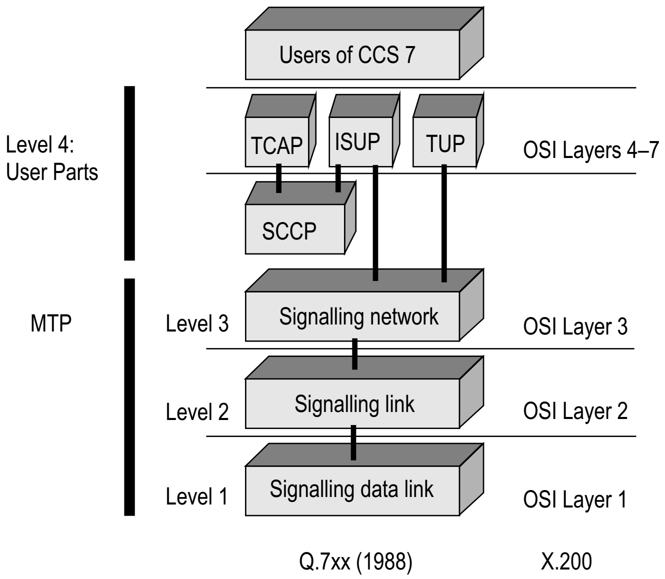


Fig. 1: The SS7 protocol stack from the OSI perspective

Telephony signalling networks have been specifically designed to – and implicitly constrained to – operate and manage 64 Kb/s circuit switched channels. There is now a strong tendency to move towards the use of packet switched networks. It is interesting to observe how packet switching, introduced in legacy networks to add flexibility and reliability to signalling in the control plane of protocol stacks, is being extended today to the user plane to carry packetized voice, merging again control and voice traffic.

2 Videoconference over packet Networks: H.323

ITU-T was the first standardization body which created an standard for transferring multimedia traffic over IP networks. The standard H.323 version 1 appeared on 1996 and was called "Visual telephony systems and terminal equipment for local area networks which provide a non-guaranteed quality of service". As a result, H.323 started the way to provide multimedia –and therefore voice transmission– over packet networks. The main contribution of H.323 was the provision of signalling protocols for controlling all of these communications, as media transmission architecture (voice, video, data) was adopted from previous works of IETF (mainly RTP/RTCP [RFC 1889] protocols tested on MBONE initiative).

After this initial version, on 1998 H.323v2 appeared with a new name, "Packet based multimedia communications systems", which have remained until now (version 4 was approved on Nov. 2000 [ITU 00a]). H.323 is considered an umbrella of standards and consists of 4 types of functional elements (Figure 2):

Figure 2: H.323 Architectural Model

- *H.323 Terminal*, is an endpoint on the network which provides for real-time, two-way communications with another H.323 terminal, Gateway, or Multipoint Control Unit. A terminal may provide speech only, speech and data, speech and video, or speech, data and video. The block functional structure of an H.323 terminal is shown in the Figure 3.

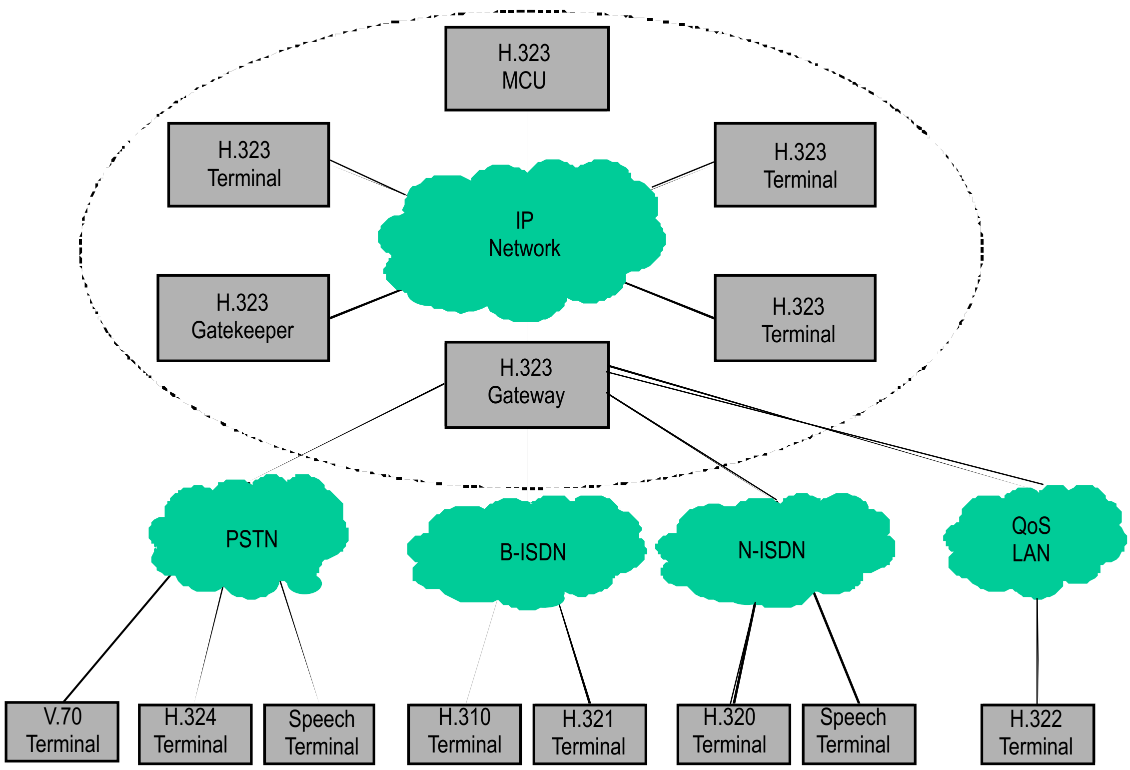


Fig. 2: H.323 Architectural Model

to Contents, page 1

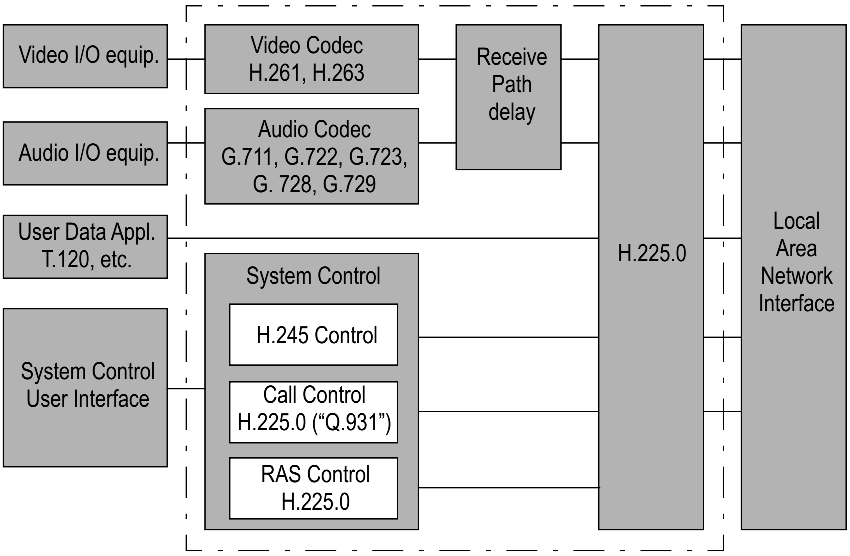


Fig. 3: H.323 Terminal structure

- *H.323 Gateway (GW)*, is an endpoint on the H.323 network which provides for real time, two-way communications between H.323 Terminals and other ITU Terminals connected, for example, to PSTN, ISDN, broadband ISDN (ATM), or QoS enhanced LANs. The purpose of the Gateway is to reflect the characteristics of a network endpoint to a Switch Circuits Network (SCN) endpoint, and the reverse, in a transparent way. Functions like translation between transmission formats and between communications procedures must be provided.
- *Multipoint Control Unit (MCU)*, is an endpoint on the network which provides the capability for three or more terminals and Gateways to participate in a multipoint conference. It may also connect two terminals in a point-to-point conference which may be later developed into a multipoint conference. The MCU consists of two parts: a mandatory Multipoint Controller (MC) which provides capability negotiation between all terminals to achieve common levels of communications and optional Multipoint Processors (MP), which performs mixing or switching of audio, video and data. This functionality could be integrated in an H.323 terminal.
- *Gatekeeper (GK)*, is an H.323 entity on the network that provides call control services to the H.323 endpoints. This element is the key block of the H.323 architecture for development of services and for the application of this technology in a medium-large environment. In principle the GK is an optional block of the architecture that was decided in order to boost the development of H.323 terminals without the requirement of a complex element, but without this element the model is quite limited. GK provides important services like: address translation and directory services, admission and bandwidth control, call control signalling or call authorization.

The complete protocol stack used in H.323 is shown in the Figure 4, including media and signalling transmission planes. Most of the control channels use TCP as transport protocol, but from version 3 on, UDP can also be used.

H.323 entities establish connections in different phases. Considering a normal scenario, an H.323 session between two entities in the same zone (that is, depending on the same GK), will involve the following steps (Figure 5):

1. *Phase A: Call Setup***.** The Calling entity uses RAS (Registration Admission and Status) protocol to send an Admission Request Message (ARQ) to the GK, requesting authorization to make a call and providing the identification of the called party. The GK can accept the call by sending a confirmation (ACF message) or reject it (ARJ message). When the call is accepted, the calling party establishes a new TCP connection to the address provided by the GK in the ACF message. This connection uses H.225.0 protocol to send a Setup Q.931 message. The Called party must first connect to the GK and ask for permission to accept the connection. When granted, the called party sends a Connect message, including the H.245 Control Channel Transport Address for use in the next phase.
2. *Phase B: Initial communication and capability exchange (H.245)*. In this phase, a new TCP connection is established between calling and called party to exchange ability and information about media channels using H.245 protocol. After this step, both parties have agreed on the media parameters (codecs, samples per frame, etc.) and exchanged information about media channels (ports, etc.). This TCP connection must remain until call termination phase and is used to open or close media channels or modify their parameters.
3. *Phase C: Establishment of audiovisual communication*. In this step, both entities exchange multimedia information directly using a RTP/UDP/IP for the media channels. They

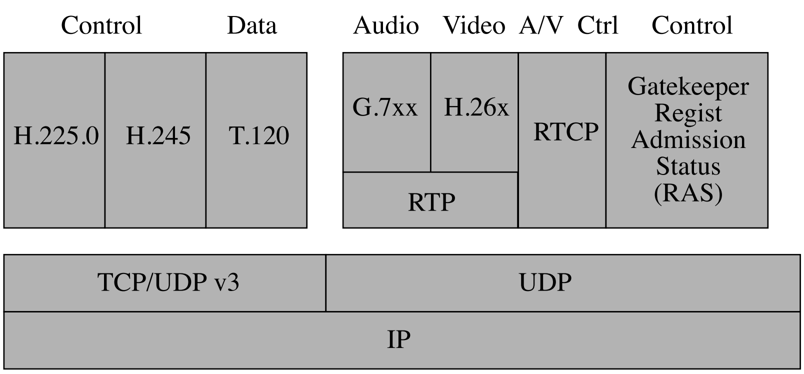


Fig 4: H.323 protocol stack

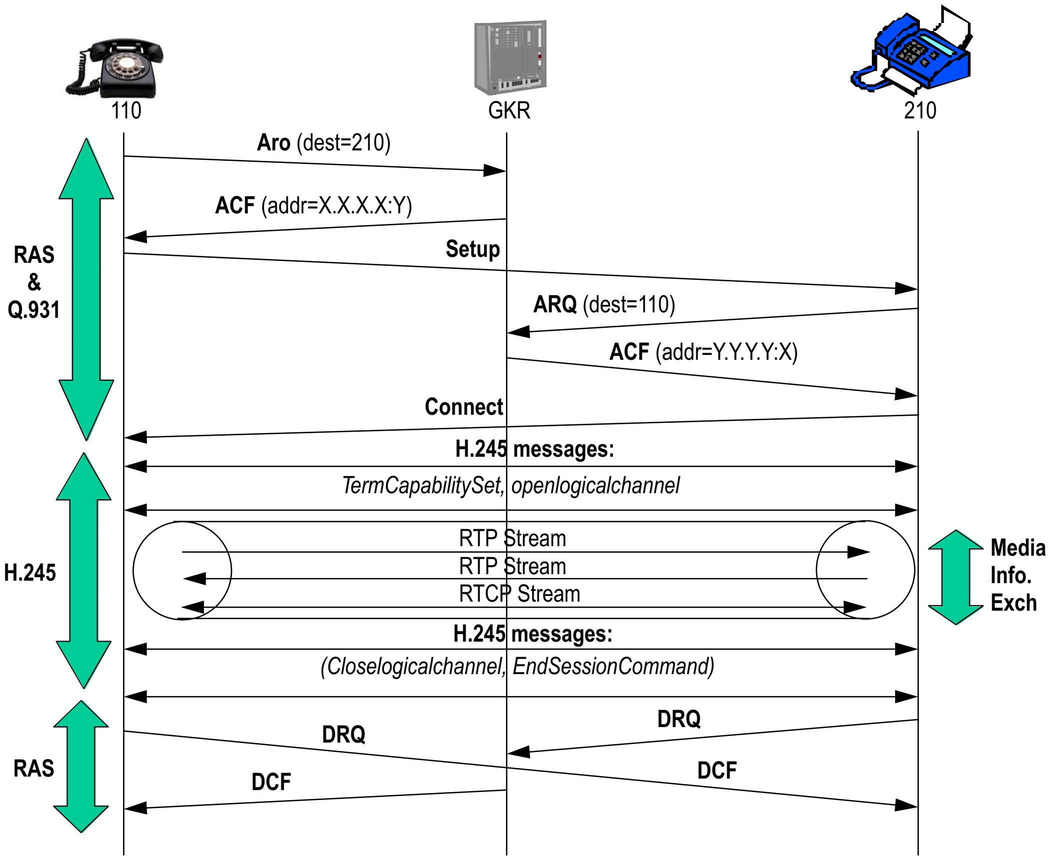


Fig. 5: Generic Call Flow

also use RTCP messages as a control channel to get feedback information about how their flows are being received.

4. *Phase D: Call termination*. H.323 entities must inform each other about the call termination by sending an *EndSessionCommand* message, which will produce the closing of H.245 channel. Besides, they must send a *Disengage Request* (DRQ) to the GK, by using RAS channel, in order to inform the GK about the termination of the call. GK then performs actions such as release resources, provide billing information, etc.

The previous example shows that H.323 requires a lot of connections before terminals can exchange information. This is one of the main drawbacks of H.323v1 that was partially solved in version 2 using two possible optional modes: Fast Connect Procedure, which allow to open media channels during H.225.0 phase; and H.245 tunnelling which uses the same channel for transmitting H.225.0 and H.245 messages.

3 IETF's proposal for VoIP: SIP

The Session Initiation Protocol (SIP) is an application protocol, defined in RFC2543 [RFC 2543], that is being designed by the IETF MMUSIC (Multiparty Multimedia Session Control) working group to enable users to participate in multimedia sessions, that is, to establish, modify and terminate multimedia sessions calls. MMUSIC working group [MMUSIC] focus on loosely coupled conferences as they exists today on the MBONE. One of the main issues in this area is related with how to inform users about forthcoming sessions, media requirements, addresses, etc. There are two basic ways to locate and to participate in a multimedia session:

- *Advertisement*. Sessions are advertised in various ways like e-mail, web pages, newsgroups or a multicast advertisements via Sessions Announcement Protocol (SAP) like in the MBONE.
- *Invitation*. Users are invited by others to participate by using the Session Initiation Protocol (SIP).

SIP has been proposed as a generic unicast and multicast initiation protocol and also as an IP Telephony call set-up protocol. It is based on a client-server protocol. SIP clients send a *Request Message* for a service, and a server handles the request, answering with a *Response Message*. SIP terminals can both generate and receive request as they are composed by a User Agent Client (UAC) and a User Agent Server (UAS).

SIP terminals can establish voice calls directly without requiring any other element. Figure 6 shows an example in which user1 calls user2 by sending an INVITE Request primitive containing user1 supported capabilities for receiving audio and a UDP port (port 12345 and mlaw codec). When user2 receives INIVITE Request, he can establish a voice channel to 12345 UDP port of user1 while he accepts the request by sending an OK Response message. In addition, user2 response includes its own media capabilities, which are used by user1 to establish a voice channel (GSM codec at 54321 port in the example) and send an ACK message to acknowledge user2's response. In order to terminate the connection, any of the

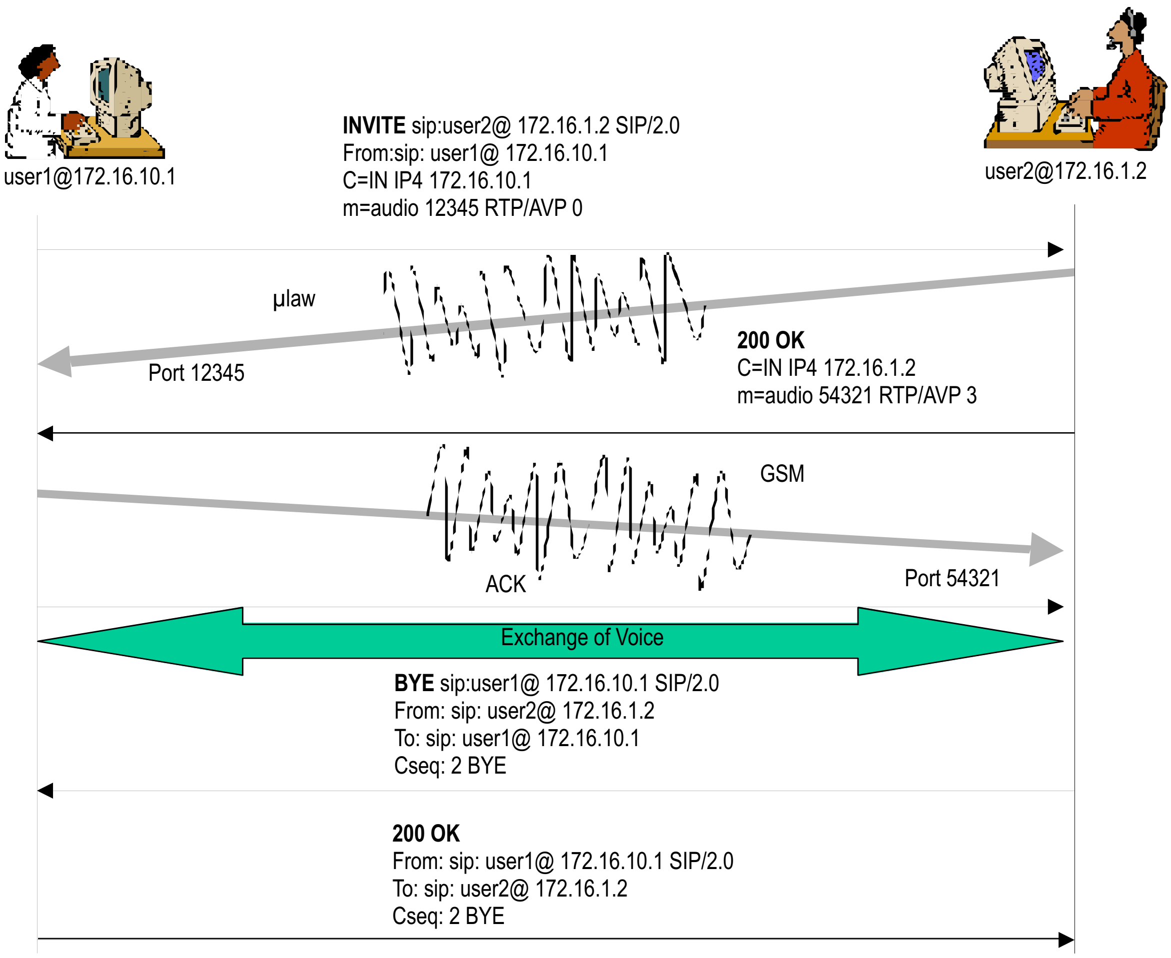


Fig 6:
Simple SIP call

parties can send a BYE Message, which must be acknowledged by an OK response.

SIP messages are encoded using HTTP/1.1 message syntax [RFC 2068], and the content of each message follows session description protocol (SDP) [RFC 2327], heavily used in the context of MBONE for distributing information about MBONE sessions.

In addition to SIP terminals which represent an IP telephones or Gateways, SIP architecture is based on four different servers entities:

- *Proxy server,* which forwards requests to its final destination. Like the Gatekeeper in H.323 model, it receives requests and sends them to the appropriate destination. In this case the Via field in the request/response message indicates the intermediate proxies which participate in the request process. This avoids routing loops as well as permits to force the responses to follow the same route back. Proxy servers do not need to relay the media except in case of transcoding operations (they only need to relay control information).
- *Redirect Server,* which, on the contrary to proxies, does not forward INVITE requests but responds to clients with a redirection message that informs about the next hop.
- *Registrar server,* that keeps track of the current location of a user by accepting registration request messages. Registrar server are discovered by using well-known multicast addresses or preconfigured unicast addresses.
- ***Call Agent***, which is a service that handles incoming and outgoing calls on behalf of a user. It is a mixture of a proxy, a registrar and a redirect server. A call agent can perform tasks like:
  - Try to find a user by redirecting the call setup message to the proper location or locations.
  - Implement call redirection rules such as call forwarding on busy, call forward on no answer, etc.
  - Implement call filtering with origin/time-dependant rules.
  - Record unsuccessful call attempts for future reference.
  - Perform any other call management task on behalf of the user.

The objects addressed by SIP are users at hosts or servers identified by Uniform Resource Identifiers [RFC 2396]. The SIP URI has the form: “user@host”, where “user” is a user name or a telephone number and “host” is a domain name or a numeric network address.

Figure 7 shows a complete example of an interaction between SIP servers. David from company.es wants to call jmoreno@upm.es. So, he sends a request to his SIP server (sip.company.es), which acts as a proxy and relies the *INVITE* request to the SIP server of upm.es (sip.upm.es). This server,

to Contents, page 1

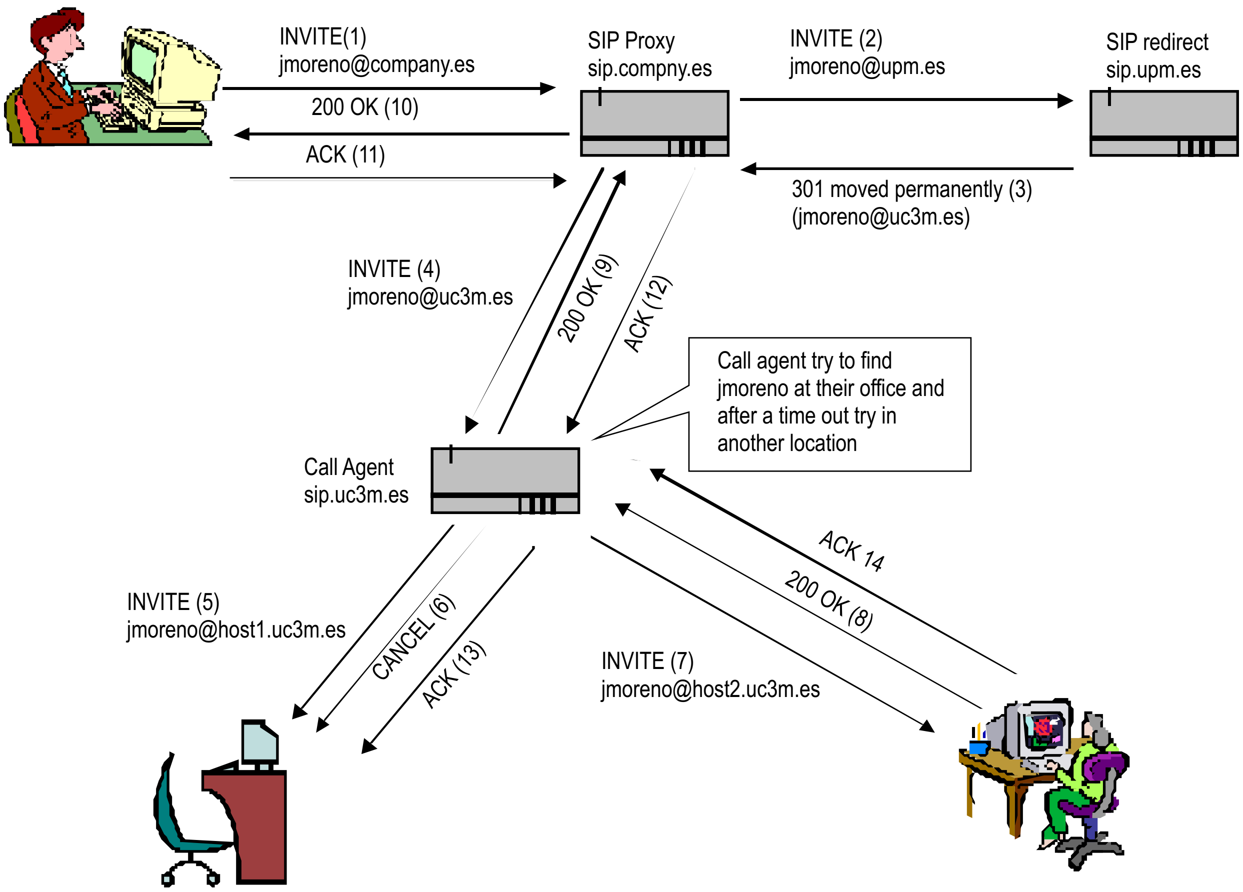


Fig. 7: Example of SIP servers

which acts as a redirect server, redirects the call by answering with a moved permanent pointing to the new address jmoreno@uc3m.es. SIP server at company.es progresses the call contacting SIP server at uc3m.es. This server, which act as Call agent, relays the request to the usual location of jmoreno at host1.uc3m.es, but after a timeout without answer from jmoreno, *CANCELs* the previous request and redirects the call to another usual location of this user in a different branch office (host2.uc3m.es). The user answers the call by sending a *200 response* which is sent back to the calling party.

One of the main advantages of SIP is its simplicity, which translates into simpler implementations and shorter delays. While H.323v1 needs 4 or 5 round-trips to establish a connexion, SIP requires only one. As mentioned before, this key aspect was corrected in H.323 from version 2 on.

4 VoIP in the core: MEGACO and MGCP

H.323 and SIP were developed bearing in mind terminals connected directly to IP networks. Both of them consider connections between IP terminals or between IP terminals and conventional SCN terminals by means of Gateways. The goal of MEGACO architecture is, in addition, to establish connections between SCN terminals through IP-based networks. The main idea comes from operators, especially incomers: What about deploying IP-based core infrastructure for new investments in the network while providing the same telephone service?

MEGACO solves this problem mainly by splitting Gateways into three different entities:

- *Media Gateway Controller (MGC)*, which provides the H.323 or SIP signalling and performs mapping between SCN signalling protocols and IP-based signalling protocols.
- *Media Gateway (MC)*, which provides media mapping and transcoding functions. It terminates SCN (PCM signal typically) and packet media signals and performs address translation, echo cancellation, playing announcements, receiving/sending DTMF digits, etc.
- *Signalling Gateway (SG)*, which provides signalling mediation between IP and SCN domains.

In a common scenario, these three elements are intended to be physically separated, providing advantages like, concentration of many MG in a few MGC controlled by a SG. Figure 1 shows MEGACO architecture.

Media Gateway Control Protocol (MGCP) provides a simple client/server model to control Media Gateways. MGCP is the result of previous protocols and has been proposed to different standardization groups like MEGACO IETF working group [RFC 2705] [RFC 30115] and ITU-T [ITU 00b]. MGCP uses SDP to exchange parameters to the gateway (IP, UDP port, media, etc.).

to Contents, page 1

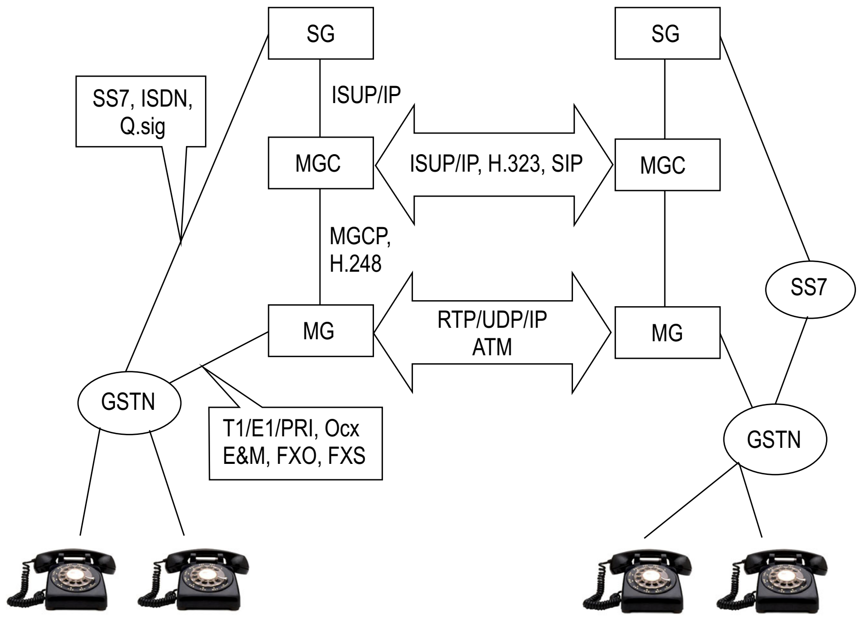


Fig 8: Megaco Architecture

5 Third Generation Mobile Networks: Towards and All-IP architecture

The Third Generation Partnership Project (3GPP) [3GPP 00a] works in the standardization of a 3G mobile system based on an evolved GSM core network and WCDMA radio access technologies. This system is the Universal Mobile Telecommunication System (UMTS) that will be launched to the market in the near future.

The first phase of the specification of UMTS was finished at the beginning of year 2000 and was called Release 1999 (R99). The 3GPP continues developing specifications to set the path of the evolution of UMTS systems. Release 4 and Release 5 (with target dates as of December 2001) are the following stages on this evolution. In this section we briefly describe the UMTS R99 architecture and the planned evolution for this architecture towards the inclusion of VoIP protocols in it [3GPP 00b].

As shown in Figure 9, the UMTS R99 architecture [3GPP 00c is basically a GSM/GPRS network [Bettstetter et al. 99 with a new access network named UTRAN (UMTS Terrestrial Radio Access Network). UTRAN is made of Radio Network Controllers (RNC) and B Nodes, that can coexist with classical GSM access network nodes (BTS and BSC). The architecture is clearly divided in two parts: *Circuit Switched (CS)* domain, used to carry voice traffic (made of MSC and GMSC nodes); and *Packet Switched (PS)* domain, used to transport data traffic

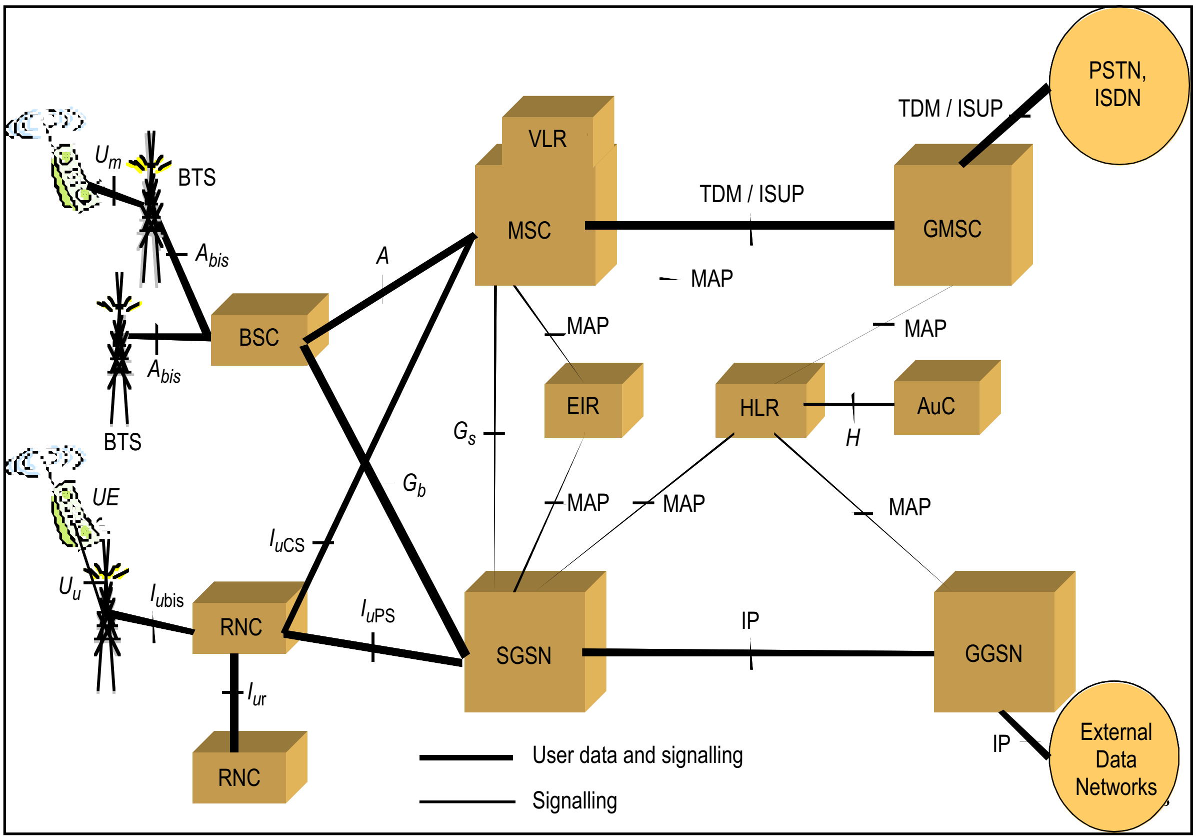


Fig. 9: UMTS R99 Architecture

(made of SGSN and GGSN nodes). Other nodes in the architecture are used to keep information about users (VLR, HLR, EIR, AuC).

The main advantage of this design is that allows an easy evolution towards UMTS systems starting from existing GSM/GPRS networks. However, there is a clear tendency to unify both domains using an IP based core network, as Release 4 and 5 point out. Even, IP could be used as the packet switching technology in the access network – not only in the core – as different European projects are proposing [MobyDick]. All this proposals have been globally named the “all-IP architecture” [Bos/Leroy 01]. The reason behind this tendency is that packet switching networks are cheaper, efficient and capable of carrying all the different classes of traffic. Besides, IP is a proved protocol that allows a seamless intercommunication with Internet.

Release 4 and 5 of UMTS specifications define the evolution steps towards the “all-IP” architecture. Key points on this path are:

- The CS domain, which was originally designed to use “classical” TDM technologies, is evolved to make it independent of the transport technology, allowing it to work over IP and ATM core networks. This evolution is highly based on the ideas developed in MEGACO architecture. For example, MSC nodes are divided in two elements, the MSC server responsible for all signalling and control functions, and the Media Gateway (MGW) responsible of data transport functions. It also uses other VoIP protocols like H.245 or RTP.
- A new subsystem, named “IP multimedia” (IM), is added to the PS domain, in order to support IP multimedia services based on SIP. The central entity in IM is the Call State Control Function (CSCF), which is in charge of call set-up and termination, routing of incoming calls, address handling, etc. It basically includes much of the functionality found on SIP servers (proxies, registrars, etc.). One of the main contributions of IM subsystem will be the use of “SIP phones”, that is, mobile terminals that directly support SIP signalling.

6 Conclusions

Signalling in telephone networks is clearly evolving from circuit switch SS7 based networks to IP centric solutions. A rich set of standardized and proprietary solutions have appeared in recent years to carry voice traffic over IP networks and cope with problems like addressing, admission and call control, internetworking with existing networks, etc. Work has concentrated over two main scenarios: voice terminals directly connected through IP networks, and operators that use IP backbones to connect traditional switched circuit networks. The first scenario can be solved by means of H.323 or SIP proposals, and the later by MEGACO or H.248.

Today, different operators and enterprises are using or beginning to use these technologies to carry voice traffic over packet networks. The trend will be increased in the near future with the evolution of mobile UMTS networks towards an “all-IP” technology, in which multimedia services will be deployed. As a result, in the future voice traffic will be mainly carried over IP technology. However, the path towards “all-IP” networks will not be easy, there is still a lot of problems to be solved, most of them related to interoperability with existing networks, and it will take a long time, investment in traditional networks has to be recovered.

7. References

[SS7 89]
Specifications of Signalling System Nº 7. CCITT Blue Book, fascicle VI.7, recommendations Q.701-Q.716, Q.721-Q.766, Q.771-Q.795. ITU, 1989.

[RFC 1889]
RFC 1889. H. Shulzrinne, S. Castner, R. Frederick, V. Jacobson. RTP: A transport protocol for real time protocol.

[ITU 00a]
ITU-T Recommendation H.323: “Packet-based Multimedia Communications Systems”, November 2000

[RFC 2543]
RFC 2543. M. Handley, H. Shulzrinne, E. Schooler, E. Rosenberg. SIP: Session Initiation Protocol.

[MMUSIC]
MMUSIC web site. http://www.ietf.org/mmusic

[RFC 2068]
RFC 2068. R. Fielding and others. Hypertext Transfer Protocol -– HTTP/1.1

[RFC 2327]
RFC 2327. 2327 M. Handley, V. Jacobson. SDP: Session Description Protocol.

[RFC 2396]
RFC 2396. T. Berners-Lee, R. Fielding, Uniform Resource Identifiers (URI): generic syntax.

[RFC 2705]
RFC 2705. M. Arango et al., “Media Gateway Control Protocol (MGCP)”.

[RFC 30115]
RFC 3015. F. Cuervo, N. Greene, A. Rayhan et al., “Megaco Protocol Version 1.0”

[ITU 00b]
ITU-T Recommendation H.248: “Gateway Control Protocol”, June 2000

[3GPP 00a]
3GPP web site: http://www.3gpp.org.

[3GPP 00b]
3GPP Technical Specification TS 23.002, v5.0.0: Network Architecture (Release 5). October, 2000.

[3GPP 00c
3GPP Technical Specification TS 23.002, v3.3.0: Network Architecture (Release 1999). March, 2000.

[Bettstetter et al. 99
C. Bettstetter, H.-J. Vögel, J. Eberspächer; GSM Phase 2+, General Packet Radio Service GPRS: Architecture, Protocols, and Air Interface; IEEE Communications Surveys Vol. 2, No. 3, 1999.

[MobyDick]
MobyDick Project. http://www-int.berkom.de/mobydick

[Bos/Leroy 01]
Lieve Bos, Suresh Leroy; Toward an All-IP-Based UMTS System Architecture; IEEE Network, Vol. 15, No. 1; 2001.

to Contents, page 1

Naming and Addressing in Voice over IP Networks

David Fernández, John Michael Walker, José A. G. Cabrera, Juan Carlos Dueñas

This paper describes how naming and addressing is being organized in Voice over IP technologies and scenarios. After describing naming and addressing schemes used in different protocols, we concentrate on specific problems found in complex VoIP scenarios. Later, we present the ENUM proposal to populate E.164 numbers into the Internet DNS system, as a solution to manage the resolution between names and addresses. Finally, we briefly analyse the role of administrations as regulators of the ENUM technology.

Keywords: ENUM, VoIP, E.164, DNS, Naming, Addressing

1 Introduction

Naming and addressing is an important function of communication networks. All entities in a network must be uniquely identified to allow data to be directed to them. In today's networks there is a clear distinction between Names and Addresses:

- A *Name* uniquely identifies an end-user or any other entity (e.g., an application) that can be communicated to via the network. Names are typically designed to be used by persons and they are composed of a combination of symbols (characters or numbers) organized in a way that can be easily remembered (e.g. hierarchical names). Usually they do not indicate the network or location of the named entity.
- An *Address* is a string or combination of digits and symbols that identifies the specific termination points of a connection or session and is used for routing. An address specifies the location of the entity in terms of network structure. It typically includes the identity of the network it is connected to and some information about the location within that network.

Naming and addressing takes place at different layers in communications architectures. For example, a typical IP host connected to a LAN has a subnetwork address (MAC), an IP address, an IP name associated with that address and several names and addresses used for the different applications and users using it. There are complex relationships between all these names and addresses and resolution functions have to be used, either locally or by means of external directories, to associate names with addresses, names with other names, or even addresses with names.

This paper deals with how naming and addressing is organized in VoIP proposals. It is organized as follows: first, we present the naming and addressing schemes found in common computer networks and applications. Then we discuss the need for resolution capabilities as a way to cope with heterogeneous network scenarios and we present DNS and ENUM as a central proposal to be used in VoIP scenarios. Later, we give some ideas about how portability can be achieved inside ENUM. Finally, some thoughts about the role of administrations in ENUM are outlined before coming to the conclusions.

2 Naming and Addressing schemes

Traditionally, telephone networks have not made a distinction between addresses and names. They have used numbers, based on the ITU-T E.164 recommendation [ITU 97], as addresses to access terminals. Assignation of E.164 numbers is made hierarchically: ITU-T assigns each country a unique country code (CC), and the naming authority in each country is in charge of delegating the rest of the numbers and assigning codes to geographic regions or different operators.

Typically, E.164 numbers have been related precisely to network architecture. In this context, they were indistinguishably

David Fernández is an Assistant Professor at the Department of Telematics Engineering of *Universidad Politécnica de Madrid*, since 1995. He received a Ph.D. degree on Telematics Engineering from that university in 1993. His current research focuses on advanced Internet protocols. <david@dit.upm.es>

John Michael Walker works as a Systems Engineer for GSM and UMTS at Ericsson. He is a member of the Numbering, Naming and Addressing (NNA) team in the System Management for Third Generation Data Bases department at Ericsson (R&D) in Madrid. He holds a Telecommunication Engineering degree from *Universidad Politécnica de Madrid*. <john-michael.walker-pina@ece.ericsson.se>

José Angel García Cabrera works at the *Comisión del Mercado de las Telecomunicaciones* as Deputy Director of the Technical Department since 1997. He is member of the ECTRA Project Team on Numbering since 1996, member of the *Cuerpo Superior de Sistemas y Tecnologías de la Información de la Administración del Estado* since 1995. He received a B.Sc. degree in Telecommunications Engineering from *Universidad del País Vasco* in 1991. <jagcabrera@cmt.es>

Juan Carlos Dueñas is an Assistant Professor at the Department of Telematics Engineering of *Universidad Politécnica de Madrid*, since 1998. He holds a doctorate degree on Telematics from the same university. His current research is on software technology for telecommunication systems. <jcduenas@dit.upm.es>

used as names and addresses. But the liberalisation that has significantly increased the number of operators, and the implementation of new functionalities like number portability, have changed this fact. Nowadays, there is a clear distinction between names, used by customers to make calls, and addresses (or routing numbers), used by the network to route calls. However, both of them are still based on the E.164 specification. As we will mention in section 4, the mapping between names and addresses is normally done using a database.

On the other hand, in IP networks there is a clear distinction between names and addresses. Names take the well known hierarchical form of domain names separated by dots, e.g. greco.dit.upm.es. Addresses are 32 bits long (or 128 bits in the case of the new version of the IP protocol, IPv6) and identify an interface of a host connected to an IP network (multihomed hosts or routers have as many addresses as they have interfaces). The association between names and addresses is at IP level and is resolved using the Domain Name System (DNS) application, which is a world wide distributed directory.

Taking a look at naming and addressing at the application layer, we see different schemes. Basic applications like telnet (remote login) or ftp (file transfers) directly use IP names, which are translated to IP addresses by the DNS. As these kind of applications are basically end-to-end (with no intermediate nodes) and sessions are normally established with machines (and not with users), the solution is quite simple.

Electronic mail uses a more elaborated scheme. E-mail names have the form “user@host”, where “host” is a domain name, which is translated to the address or addresses of the hosts that provide e-mail services to the user being named using the DNS (and the MX records stored in it).

WWW naming is based on the well known Uniform Resource Locator (URL) names. URLs, which have been generalised and named Uniform Resource Identifiers (URI) [Berners-Lee et al. 98], are becoming the standard way of naming resources on the Internet.

That is the case, for example, of SIP [Hersent et al. 00], one of the main VoIP signalling proposals. SIP uses SIP-URLs to identify “users at hosts”. They take a form similar to mailto URLs (e.g., mailto:user@domain), for example:

- sip:david@dit.upm.es, which is the simplest form of SIP-URL and identifies a user (david) in a domain (dit.upm.es),
- sip:david:secret@dit.upm.es, which is the same as the previous example but includes a password,
- sip:+1-212-555-1212:1234@gateway.com;user=phone, which identifies a telephone number which is accessible through a gateway located at the domain gateway.com (1234 specifies a password).

There are also other naming schemes based on URL formats being used in VoIP. The most common one is the *tel-URL* scheme [Vaha-Sipila 00], that allows the specification of E.164 numbers inside a URL (e.g., tel:+358-555-1234567). There are also formats to specify fax and modem numbers.

Finally, in H.323 [Hersent et al. 00], calls are directed to hosts where H.323 applications reside. That is why H.323 bases its naming and addressing scheme on IP names and addresses. However, in order to add flexibility, H.323 includes a complex system of aliases that allows users to make calls using, for example, e-mail addresses, E.164 numbers or even the new H.323 URLs defined in the latest version of H.323. The Gatekeeper, which is the entity that controls H.323 systems in a zone, is in charge of translating the aliases into the network address for the destination terminal, using its internal translation tables or external directories like DNS.

3 Resolution in VoIP Networks

Each naming and addressing scheme defines its own rules to make resolutions possible, that is, to define the functions that allow the conversion between names and addresses at the same or different communication layers. Resolution can be carried out in very different ways depending on the scenario. In some cases, a simple table can solve the problem. That could be the case, for example, for a small VoIP network based on H.323 with only one Gatekeeper that centralises the information about terminals registered in its zone. Terminals will just ask the Gatekeeper when they need to resolve a name or address.

But in general, VoIP scenarios are much more complex than the one described above for several reasons:

- Scalability: VoIP solutions are being designed to scale to the size of present telephone networks or even larger,
- Reliability: VoIP solutions should reach reliability figures similar to present telephone network switches (“five nines”, that is, 99.999%),
- Interoperability: VoIP scenarios are not going to be homogeneous; interoperability between VoIP networks based on different solutions (e.g. SIP and H.323) and, of course, with the present telephone networks have to be guaranteed.

All these facts, and particularly the last one, raise the complex problem of having to harmonize all of the naming and addressing plans involved in VoIP scenarios. To study the problem in detail, TIPHON (Telecommunications and Internet Protocol Harmonization Over Networks), a project inside ETSI, has defined five different scenarios for a telephone call [ETSI 99]:

- Scenario 0: a call between two IP terminals,
- Scenario 1: a call from an IP terminal to a Circuit Switched (CS) terminal that goes through a gateway,
- Scenario 2: a call from an CS terminal to an IP terminal through a gateway,
- Scenario 3: a call between two CS terminals that go through an IP backbone (traversing two gateways), and
- Scenario 4: a call between two IP terminals that go through a CS network (traversing two gateways).

For example, Figure 1 presents the steps to be carried out in scenario 2. As shown, several resolutions have to be made to complete a telephone call. First, an initial resolution could be made by users to get the number (strictly speaking, the name) of the person they want to call. Then a Service Resolution has to be carried out in order to get the address of the network where the destination is. Service Resolution could be used to support number portability or non-geographic services such as freephone. Finally, Routing Resolution has to be invoked, maybe several times, to route the call to the destination.

At present, several proposals are being discussed for performing this resolution on a large scale. Some of these propose

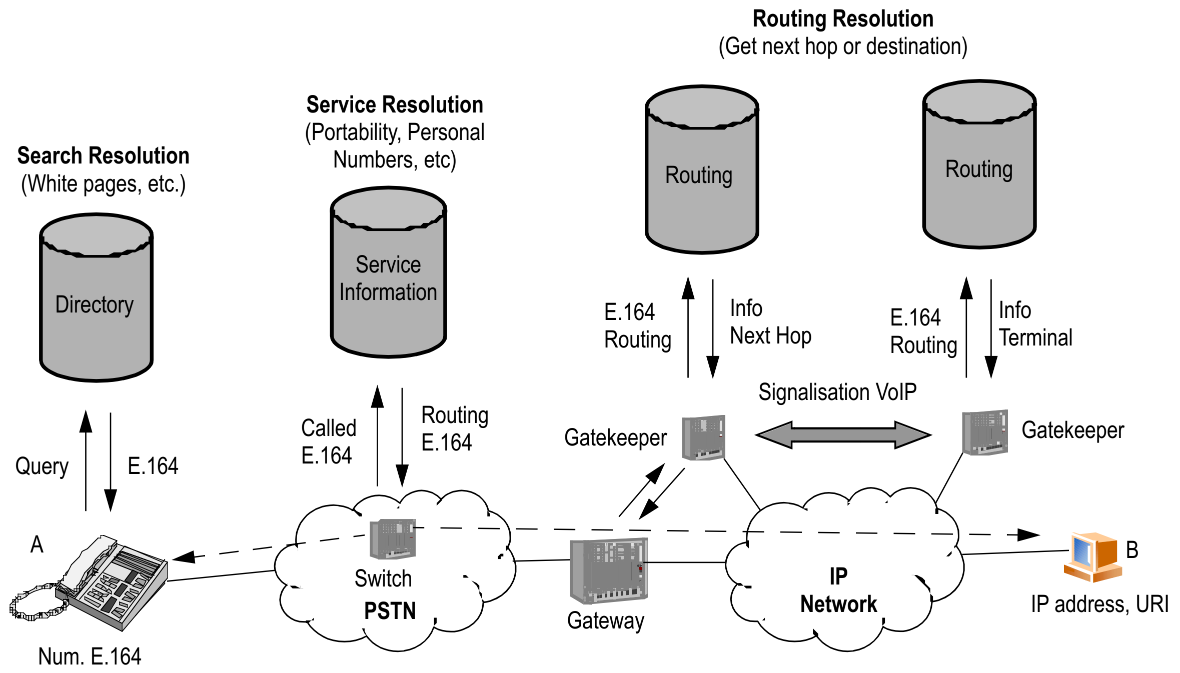


Fig. 1: Resolutions in TIPHON Scenario 2

reusing the solutions, protocols and experience gained from the Internet using the Domain Name System (DNS). DNS is a world wide hierarchical distributed database that provides the capability of discovering the IP address of a certain network element, once its name is given. DNS is also able to carry out other types of resolutions, for example, from mail domains to mail server addresses or from IP addresses to names (inverse resolution).

ENUM is the main initiative in this field. It basically proposes including E.164 numbers in the DNS information tree and using them as the key to look for information about the services associated with a number. ENUM is described in detail in the following subsections.

3.1 E.164 Numbers and the Domain Name System (DNS)

E.164 numbers identify many different types of end terminals, supporting many different services and protocols: ordinary phones, fax machines, pagers, data modems, e-mail clients, text terminals for the hearing impaired, etc. This leads to the following problem: how can users discover services and protocols supported by each E.164 number?

A working group named ENUM was set up in IETF to address the above challenge via DNS-based architecture and protocols for mapping a telephone number to a set of attributes (e.g. URIs) that can be used to contact a resource associated with that number. The result of this working group has been RFC 2916 [Faltstrom 00] that proposes the use of DNS to store E.164 numbers and to identify services linked to it.

The actual transformation of E.164 numbers into DNS names is realised as outlined in the following example. Take an E.164 number written in its full form, including the country code without any non-digit characters, except the international "+" sign, e.g. +34918761234. The "+" sign is removed and dots are placed between each digit, e.g. 3.4.9.1.8.7.6.1.2.3.4. Finally, the number is placed in reverse order and the string "e164.arpa" is appended, e.g. 4.3.2.1.6.7.8.1.9.4.3.e164.arpa

E.164 numbers populated into the DNS must comply with the hierarchical model inherent to the DNS. Hence it was agreed that they should all be populated under a specific domain "e164". The Top level domain ".arpa" was chosen; as this domain allows reverse address lookups, that is, allows the mapping of an IP address to a domain name. In a similar manner, ENUM maps a number (understood here as an address) to a URI that may be another domain name. Another reason behind choosing ".arpa" versus other possible top level domains is that the domain was redefined as an acronym for "address and routing parameters area" (previously it was short for ARPAnet). The object of this is to centralise all of the infrastructure domains (e.g. .e164) under one top level domain.

Both the DNS and E.164 numbering scheme are hierarchical and delegated structures. Figure 2 shows an example of how an E.164 number is delegated throughout different DNS servers while keeping to the E.164 hierarchy.

Each server is delegated a certain section of the number. Under .e164 a set of country code servers would cover an administrative (or sovereign) E.164 domain, e.g. ".6.4." for Sweden. The higher level country-code server then contains entries for each operator in that administrative domain and so on.

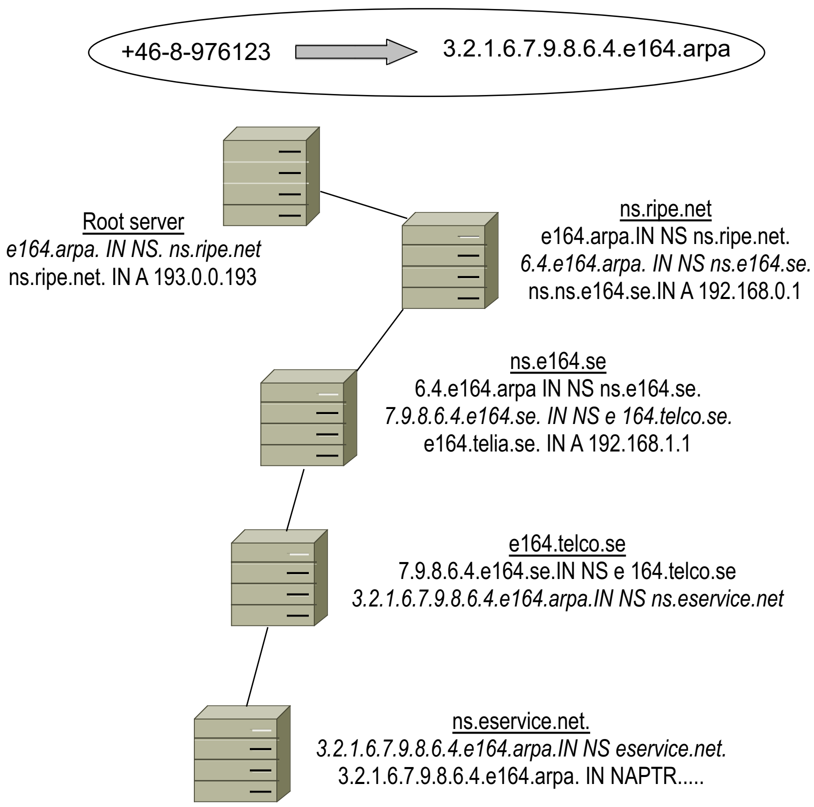


Fig. 2: Example of E.164 number delegation in the DNS

3.2 Converging Networks

ENUM is judged as one of the principal mechanisms for convergence between the Internet world and the PTSN environment. Many people are referring to ENUM as the "Internet's Glueball infrastructure[1]" due to ENUM's ability to bond an E.164 number to a range of internet-based services via URIs. ENUM is considered an internet-PSTN convergence enabler due to Naming Authority Pointers, NAPTRs, defined in RFC 2915 [Mealling/Daniel 00].

Domain names identify nodes that store resource information in what are known as resource records (RR). These RRs may store IP addresses, names of other servers, pointers to other parts of the domain name space[2], etc. The NAPTR is an RR type that includes a regular expression that can be used by a client program to rewrite a string into the domain name.

While ENUM can make use of resource records that may resolve a telephone number into an IP address, another host name, etc, its real potential lies in bonding a number to NAPTR resource records. This means that an E.164 number may be mapped into another URI. This in turn means linking a number to many new services identified by those URIs.

1. Glueballs are special kinds of particles in sub-atomic physics that are believed to hold everything together.
2. The reader is referred to RFC 1034 "Domain Names – Concepts and Facilities" [Mockapetris 87] for more on this.

The example in Figure 3 is taken from RFC 2916.

In the above example, a DNS query for number +4689761234 results in a list of NAPTR resource records. The records for this number show the following services as being available: multimedia (SIP), e-mail (mailto) and voice telephony (tel). Note that each service is a new communication medium linked to a new identifier.

The DNS client must be able to handle the multiple NAPTR records returned. The client will apply all substitutions and performs all lookups, not the DNS servers. The client uses the following key fields in the records to process the response:

- Order, which is a value that sorts multiple NAPTR records in a response.
- Preference, which indicates the preferred service from amongst those with the same order value.
- "U", which is a flag indicating that a URI must be returned.
- "E2U", which is a service that indicates an E.164 number is mapped to the URI.
- "Regular expression" which indicates the URI rewrite rule; this will replace the number with the specified URI.

In the previous example, the preferred communication method is SIP as it has highest order. The rewrite rule should replace the E.164 number with the required SIP URI. The client should be able to detect loops in case the last TEL URI is chosen.

Other standardisation forums are already applying ENUM as a mechanism for convergence between networks. In particular, 3GPP[3] have recently introduced ENUM for translation of E.164 numbers to SIP URIs in the multimedia subsystem.

$ORIGIN 4.3.2.1.6.7.9.8.6.4.e164.arpa.

	order	*pref*	*flag*	*service*	*regexp*
IN NAPTR	10	10	"u"	"sip+E2U"	"!^.*$!sip:paf@ericsson.net!".
IN NAPTR	102	10	"u"	"mailto+E2U"	"!^.*$!mailto:paf@teleco.es!".
IN NAPTR	102	10	"u"	"tel+E2U"	"!^.*$!tel:+4689761234!".

Fig. 3: Example from RFC 2916

4 Number Portability Issues and ENUM

Number portability refers to the ability of a subscriber to change service provider without a change in telephone number. E.164 number portability is addressed in recommendation ITU-T E.164 Supplement 2. In other words, ENUM must also be able to cope with portability as it can be considered an inherent property of all E.164 numbers. Another more basic

3. Third generation partnership project, www.3gpp.org

to Contents, page 1

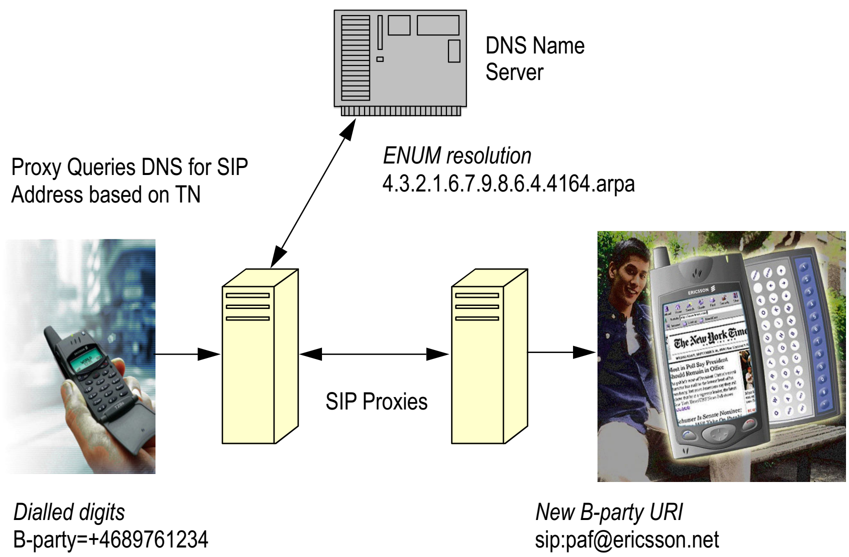


Fig. 4: ENUM as an enabler of communication between PSTN and multimedia users

reason for ENUM being able to deal with number portability is because it is a regulatory requirement.

In the PSTN environment, the number portability problem requires a means of routing towards a new service provider for a given E.164 number. This implies performing a database lookup to obtain additional routing information, returning either a routing number or a directory number.

In the DNS world, the problem is treated as a re-delegation of the E.164 entry to a new service provider. In practice, this is achieved by changing the name server records "NS" to point to the new provider. A higher level authority, presumably the registry / national-regulatory-authority running the country-code database for a sovereign E.164 region, should oversee the portability process for an ENUM based mechanism.

Another case of portability that appears with ENUM is when a specific service for a given telephone number is changed to a new provider. The NAPTR record indicating the specific service must be updated to show this change. This process would be coordinated by the ENUM service registrar (the service provider that actually links a number to a set of NAPTR records on their name server).

Actual administrative and technical problems regarding number portability and ENUM are still under consideration. Both ITU-T and IETF are proposing solutions that will ensure correct administrative procedures and allow for compatibility with number portability solutions in the PSTN environment.

5 Role of Administrations

ENUM is a key technology based on DNS and other resources from the IP networks, developed by the Internet world. Nevertheless, as the input to perform ENUM translations is an E.164 number, the IITU-T as well as the national authorities in charge of numbering plans management, plays a crucial role in the implementation of ENUM systems and services.

The outcome of several meetings amongst people from the Internet Engineering Task Force (IETF) and the ITU-T [ITU 00] has been a decision on how to focus on some very relevant issues of an administrative nature.

Although most of the questions are still open, key regulatory points that fall within the scope of the ITU itself and of the National Regulatory Authorities (NRAs) have been identified. In fact, most ENUM service and administrative decisions are national issues under the purview of NRAs, since ordinary E.164 resources are managed nationally.

At the top level, the ITU-T has the responsibility for providing E.164 assignment information to DNS administrators, for performing the administrative function. The ITU will ensure that each NRA has authorised the inclusion of their country code information for input into the DNS. Below this layer, each NRA has the entire responsibility of the regulation of ENUM and provides the E.164 assignment information to DNS administrators for performing the administrative function.

The most important challenges for regulators in order to facilitate the implementation of ENUM, and to guarantee a level playing field for the development of fair competition, are the following:

- NRAs shall decide on the way that administration and control tasks are structured in hierarchical levels. This aspect is seen as a key point as the administrative side of ENUM is considered to be more complex than it is technical.
- NRAs also have to ensure that all parties have the information needed to perform its task, taking into account the need for accuracy of the information held in the ENUM service and the rights of privacy of the end users. The NRA itself has to make available all the information related to numbering allocation and number portability.
- ENUM allows ISPs to easily provide public voice telephony, requiring E.164 numbering resources for the identification of dial-up access users. Some countries may have difficulty in accommodating growth in demand for geographic numbers.
- It is the role of each NRA to select and appoint registries for the part of the DNS that relates to their country.
- NRAs shall promote the openness of the ENUM architecture to competition, in order to prevent entry barriers to the market and other causes of distortion of competition.

The European Commission as well as European regulators, individually or together in the project teams of the European Committee for Telecommunications Regulatory Affairs (ECTRA), are studying ENUM matters with the aim of allowing the deployment of this technology under a fair and non discriminatory competitive framework.

For the time being no regulatory decisions have been taken regarding ENUM, but regulatory bodies are working in this field in order to prepare the framework for the provision of this kind of service. NRAs, according to the minimal intervention principle, have to allow the deployment of all the types of services that this technology is capable of providing, and at the same time have to ensure that the provision of such services abides by market and privacy legislation.

6 Conclusions

This article has presented an overview of how naming and addressing is being organized in VoIP networks. The complexity of the scenarios in which VoIP technology is going to be deployed, in terms of scalability, reliability or heterogeneity, poses important challenges in the solutions being designed. ENUM is a leading initiative that tries to contribute and incorporate all the experience gained in the Internet's DNS system to the solution of some of these important problems. there are, however, still some uncertainties about how the DNS will behave when managing a number of entries one order of magnitude higher than the number being supported now.

However we believe that in the near future ENUM will play an important role in VoIP networks, not only as a mechanism to be used for resolution purposes, but, what it is most important, as an enabling technology to allow the deployment of new applications, which is in fact the main reason behind all the efforts being invested in VoIP nowadays.

References

[Berners-Lee et al. 98]
T. Berners-Lee, R. Fielding, L. Masinter. "Uniform Resource Identifiers (URI): Generic Syntax". RFC 2396, August 1998.

[ETSI 99]
ETSI TS 101 324. "TIPHON; Numbering; Scenarios 1, 2, 3 and 4", November 1999.

[Faltstrom 00]
P. Faltstrom. "E.164 number and DNS". RFC 2916, September 2000.

[Hersent et al. 00]
O. Hersent, D. Gurle, J.P. Petit. "IP Telephony", Addison-Wesley, 2000.

[ITU 97]
ITU-T E.164 Recommendation. "The international public telecommunication numbering plan", 1997.

[ITU 00]
"ITU-T SG2 Liaison to IETF/ISOC on ENUM". ITU-T Working Party 1/2, Berlin, 19-26 October 2000.
http://www.itu.int/infocom/enum/wp1-39_rev1.htm

[Mealling/Daniel 00]
M. Mealling, R. Daniel "The Naming Authority Pointer (NAPTR) DNS Resource Record", RFC 2915, September 2000.

[Mockapetris 87]
P. Mockapetris, "Domain Names-Concepts and Facilities", RFC 1034, November 1987.

[Vaha-Sipila 00]
A. Vaha-Sipila. "URLs for Telephone Calls". RFC 2806, April 2000.

to Contents, page 1

Multimedia Services over the IP Multicast Network

Antonio F. Gómez-Skarmeta, Angel L. Mateo, Pedro M. Ruiz

Voice over IP (VoIP) is one of the most important and complex new services that are being introduced in Internet. VoIP makes use of several different technologies like signalling, streaming of real time data, session management, etc. The development and the experimentation of video conferencing applications over IP multicast networks have contributed greatly to the maturation of some of these technologies. This article summarizes the most important topics related with IP Multicast technology and video conferecing over IP Multicast networks. After introducing IP multicast technology as a mean to support many-to-many communications, we present some of the protocols and the applications used over IP multicast service. Finally, we outline some of the problems that preclude IP multicast to be widely deployed.

Keywords: IP Multicast, MBone, MBone Tools, RTP, SAP, SDP, Multicast Routing

1 Introduction to IP Multicast

We are presently go through a revolution in the Internet world as it happened with the WWW. Terms like e-commerce, video conferencing, streaming, video on demand and many others are becoming common usage in day-to-day life. VoIP is one of the technologies contributing to this revolution. However, the present IP networks are not adequately adapted to support this kind of services. For example, there are no mechanisms that guarantee a satisfactory quality of service (QoS) for VoIP communications. To support many-to-many communication, IP multicast offers a much more efficient mechanism than the current IP unicast networks.

1.1 Unicast vs. Multicast

The typical Internet services are based on the IP unicast model, that is to say: datagrams are addressed to only one host. Such communication is called "one-to-one". In some other situations, when there are more than two parties, the use of IP unicast can be very inefficient because the same information has to be sent to several destinations, and this process could overload the senders and the network (Figure 1).

How can we avoid this problem? Traditionally, it has been solved using "reflectors" (or Multipoint Control Units – MCU – according to H.323 terminology). A reflector is the equipment responsible for sending a packet from a source to all destinations taking place in the communication. This approach has several drawbacks, the most important being the excessive bandwidth consumption.

As an alternative, IP multicast can be used to make a datagram reach all the destinations that belong to a group. The concept of group is implemented using a special range of IP addresses. When a host is interested in receiving the datagrams addressed to a group, it has to join that group by signalling it to the network. This subscription is completely dynamic. The important fact behind IP multicast is that the source only sends one packet and the network is responsible for making the necessary copies to reach all destinations. This copying is made so that only one instance of the packet is transmitted over each link.

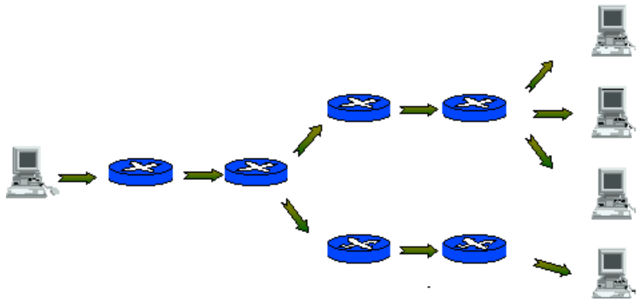


Figure 1: Distribution of one only packet to multiple receivers

Antonio F. Gómez-Skarmeta is an Assistant Professor at the Department of Communications and Information Engineering of the *Universidad de Murcia* since 1996 where he received a Ph.D. degree in Computer Science from that university. His current research focuses on multimedia Internet protocols and quality of services. <skarmeta@dif.um.es>

Angel L. Mateo is a telematic service manager at the *Universidad de Murcia*, where he is also an associate professor at the Department of Communications and Information Engineering. He is working on his Ph.D. degree and he is the responsible of video conferencing services in that university. <amateo@um.es>

Pedro M. Ruiz graduated from the *Universidad de Murcia* with a degree in Computer Science. He is currently an associate professor at the *Universidad Carlos III* of Madrid and works in the network engineering area of the Spanish National Research Network (RedIRIS). His main interests are advanced network services and protocols. <pedro.ruiz@rediris.es>

to Contents, page 1

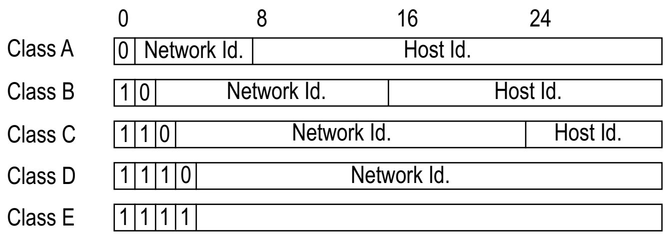


Figure 2: IP Address classification

1.2 Multicast Addressing

When the IP addressing scheme was designed, several classes of addresses were defined, as depicted in Figure 2. The class D addresses were reserved for multicast. So, IP multicast uses the range 224.0.0.0 to 239.255.255.255. These addresses are commonly called group addresses or multicast addresses. From the whole range of multicast addresses, some are reserved for specific purposes, and the rest can be used by multicast applications.

1.3 Internet Group Management Protocol (IGMP)

As we mentioned before, the way IP multicast works can be summed up as follows: receivers join the group they are interested in, and the network makes the datagrams sent to that group delivered to every receiver that joined the group. A mechanism is needed for the hosts to tell the router that they are interested in joining a certain group. This mechanism is the IGMP protocol [Deering 89], that defines the behaviour of hosts and routers when informing about joining or leaving a group.

When a host wants to join a group, it sends an IGMP report message to the 224.0.0.1 address (which should be included with every multicast-enabled host). When the router receives this message, it takes into account that there is at least one host interested in receiving that group on that interface.

Periodically, the router sends IGMP query messages to the IP multicast address of the group to ask for renewals. If there are still receivers, one of them must answer with an IGMP report addressed to the group it wants to renew. The IGMPv2 is now commonly used, however IGMPv3 implementations start coming up.

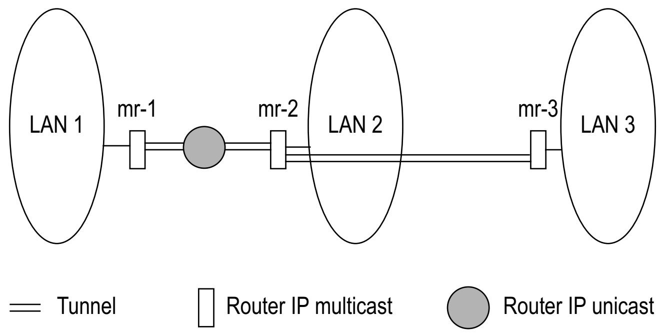


Figure 3: A multicast tunnel schema

1.4 Mrouters and tunnels

Most of the first Internet routers were manufactured without taking into account IP multicast traffic[1]. In order to experiment with IP multicast over Internet, it was necessary to define a way to interconnect IP multicast-enabled networks through networks without IP multicast support. This activities brought about what we know today as the Multicast Backbone, or simply MBONE [Macedonia/Brutzman 94].

IP multicast-enabled routers are sometimes called *mrouters*. They must satisfy two basic requirements:

- Implement the IGMP protocol.
- Use some IP multicast routing algorithm.

In order to connect a network to MBONE, one of their *mrouters* has to be connected to the rest of the IP multicast clouds. If our Internet Service Provider (ISP) offers the native IP multicast service, a simple configuration will suffice to do the task. Otherwise, we will need to configure a "tunnel" to another router connected to the MBone. This tunnel will be used to encapsulate every IP multicast datagram that has to be transmitted between these two routers into IP unicast datagrams addressed to the other end of the tunnel. So, an IP multicast-disabled network can be traversed (Figure 3).

1.5 Multicast routing

IP multicast routing is in charge of making the datagrams to flow from sources to every destination joined to a group. A good routing algorithm should guarantee that

- an IP multicast datagram addressed to a multicast group G, reaches all the hosts that have joined G,
- there are no loops, i.e., a datagram reaches its destination only once and, if possible, using the shortest path.

The IP multicast routing protocols are usually classified according to the way they work: There are protocols that work in *dense mode* like *Distance Vector Multicast Routing Protocol* (DVMRP [Waitzman et al. 98]). Some others work in *sparse mode* like *Protocol Independent Multicast Sparse Mode* (PIM-SM [Estrin et al. 98]). There are algorithms that do not fit exactly in one of these two models, an example is *Multicast Open Shortest Path First* (MOSPF [Moy 94]). The common practice is to use *PIM-sparse-dense mode* as intradomain multicast routing protocol because it is very efficient and needs no special routing messages interchanged between neighbours. Instead, it uses the unicast routing table in the router to do the calculations and decide on the better paths.

Multicast routing algorithms are usually much more complicated and difficult to configure than the unicast ones.

2 Services over IP Multicast

2.1 Multimedia Protocols

IP multicast as a network service offers no service directly to the user. But it offers an excellent framework for many-to-many multimedia data internetworking. Many protocols have come up in the IP multicast and multimedia content distribution

1. Almost every router manufactured today implements at least IGMP and one or more IP multicast routing protocols.

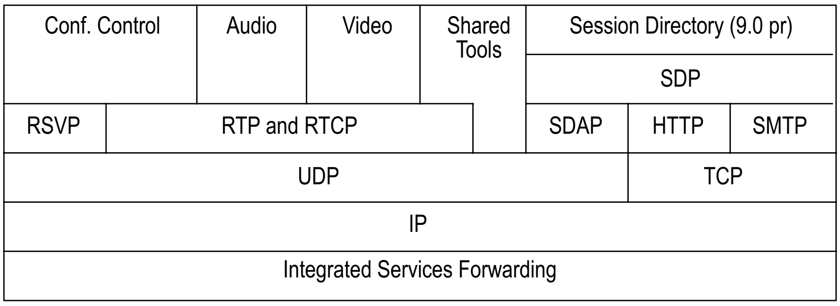


Figure 4: Multimedia protocols

environment (Figure 4). This set of protocols range from generic protocols like those used for audio and video streaming (i.e. RTP), or to manage multimedia sessions (i.e. SAP, SDP, SIP); to others much more specific used only for certain applications. It is important to note that some of these protocols, although designed, developed and tested over MBONE, have been widely adopted in other frameworks like VoIP.

2.2 Real Time Protocol (RTP)

Internet follows a *best-effort* delivery model. That means that, when a datagram is sent, the network does guarantee neither the delivery nor that the different datagrams sent arrive in correct sequence or in time. The network only guarantees that it will do as well as possible in delivering the datagram. In some cases, specially with real-time data like VoIP traffic, this model does not work properly.

Besides, when sending continuous media (like digitised voice or video) over a packet network, it is essential to provide the means for the receiver to be able to reconstruct the original information. The packets generated at the origin – typically a flow of packets equally spaced in time – will arrive to the receiver possibly out of sequence, with losses or changes in inter-packet times (jitter). A mechanism is needed for the receiver to be able to reproduce the audio or video being sent in such environment.

Real-Time Transport Protocol (RTP, [Schulzrinne et al. 96]) is the protocol defined by the IETF for real-time data transport over the Internet. The protocol basically labels each datagram to be sent, attaching them information like the time when they were generated or the type of codec being used, so that when they reach the receiver, it can reproduce the original flow adequately.

It is important to note that RTP provides no mechanism for on-time delivery or any other QoS guarantees. In order to offer QoS support for real-time communications, RTP have to be used together with other protocols or solutions like Integrated Services or Differenciated Services QoS models. However, RTP provides specific means to control the quality of the distributed data, the Real Time Control protocol (RTCP), that allows the senders and receivers to know, for example, the packet loss rate in a session. In addition, RTCP provides identification mechanisms for RTP communications.

RTP has become the *de facto* standard to send continuous media over packet (mainly IP) networks. It is used by most of VoIP proposals like H.323, MEGACO or SIP.

2.3 SAP and SDP

In a multicast environment where all participants can send and receive data between them, the concepts of client and server makes no more sense. Hence there is a need for a mechanism to allow for the user to locate a conference he is interested in. This mechanism must be based in the concept of session, which is an aggregation of related contents. For example, in a video conference, a session would be defined as the multicast group used for audio transmission, the multicast group used for audio transmission, the codecs being used both for audio and video, and so on.

For session management, two protocols have come up in MBONE: SDP and SAP. SAP (*Session Announcement Protocol* [Handley et al. 00]) defines the use of specific multicast groups to distribute session information. The session creator is responsible for periodically reannouncing it so that people joining the special "announcement group" after the creation can still know about the session. The information used to describe the session is specified by SDP (*Session Description Protocol* [Handley/Jacobson 98]).

As RTP, SDP protocol has been reused in other environments, for example, it has been incorporated into MEGACO VoIP proposal.

2.4 SIP

The SDP/SAP model requires the user to look for the session he wants to attend. However, in some scenarios like IP telephony, a way to invite other parties to participate in a session is needed. The *Session Initiation Protocol* (SIP, [Handley et al. 99]) came up for covering this need.

SIP defines the signalling mechanisms that are necessary to establish a session and to negotiate the parameters to be used in it, such as codecs, media, location, etc. As other protocols mentioned, SIP has surpassed the MBONE environment were it was originally created and now it has become one of the main proposals for VoIP. In fact, SIP has been recently selected by 3GPP to be used as the VoIP protocol for 3G mobile networks based on "All-IP" proposal.

3 MBone Tools

Several applications have been developed to test the advantages of the IP multicast model at the initial stages of the MBone. These tools, commonly known as the *MBone Tools* allow us to participate in different kinds of video conferences and meetings using IP multicast as the network technology. The typical MBone tools are (Figure 5):

- SDR. This tool is equivalent to a TV guide. It shows all planned and ongoing MBONE sessions. Recent versions also allow us to use a "quick call" service based on SIP.

to Contents, page 1

- VIC. This tool is used for video transmission with a great variety of codecs available. It can be used on almost every platform and is compatible with several standards for capturing video. So, it allows a simple personal computer to send video without needing to buy an expensive video capturing hardware.
- VAT and RAT are used for audio conferencing. They are also available for many platforms and support several codecs like GSM, PCM, DVI, and so on.
- WB. This tool is a distributed shared whiteboard that can be used by all the participants and offer the same functionality as the usual blackboard in a classroom.
- NTE stands for *Network Text Editor* and offers the functionality of a distributed word processor. It supports tokens for asking permission to write and is quite comprehensive.

Recent applications using IP multicast are more complex but offer new important functionalities. The goal is to integrate all these tools into one specific tool possibly in the Web. In fact, some big projects like MASH are working on Web integration. Some of these "new generation multicast tools" are:

- DLB is an improved version of an electronic whiteboard that supports two modes (on-line and off-line). It thus allows for editing slides off-line and then present them on-line.
- MiNT is a very complete application developed by the German GMD that supports the SIP protocol, includes an RSVP agent for bandwidth reservation and even an integrated GUI for audio and video transmission.
- MASH is a very big project initiated at Berkeley and its main goal is to integrate the typical MBone Tools into a common GUI. In addition, tools for playing and recording sessions are offered. They are also deploying "mashlets" that aim to integrate the GUI into the WWW.
- RELATE (REmote LAnguage TEaching). This tool was developed at the University College London and is very interesting for teaching on-line. Although it was initially thought for language teaching it can also be used to teach some other subjects. This tool integrates into the same GUI the audio, video, whiteboard and text editor applications.

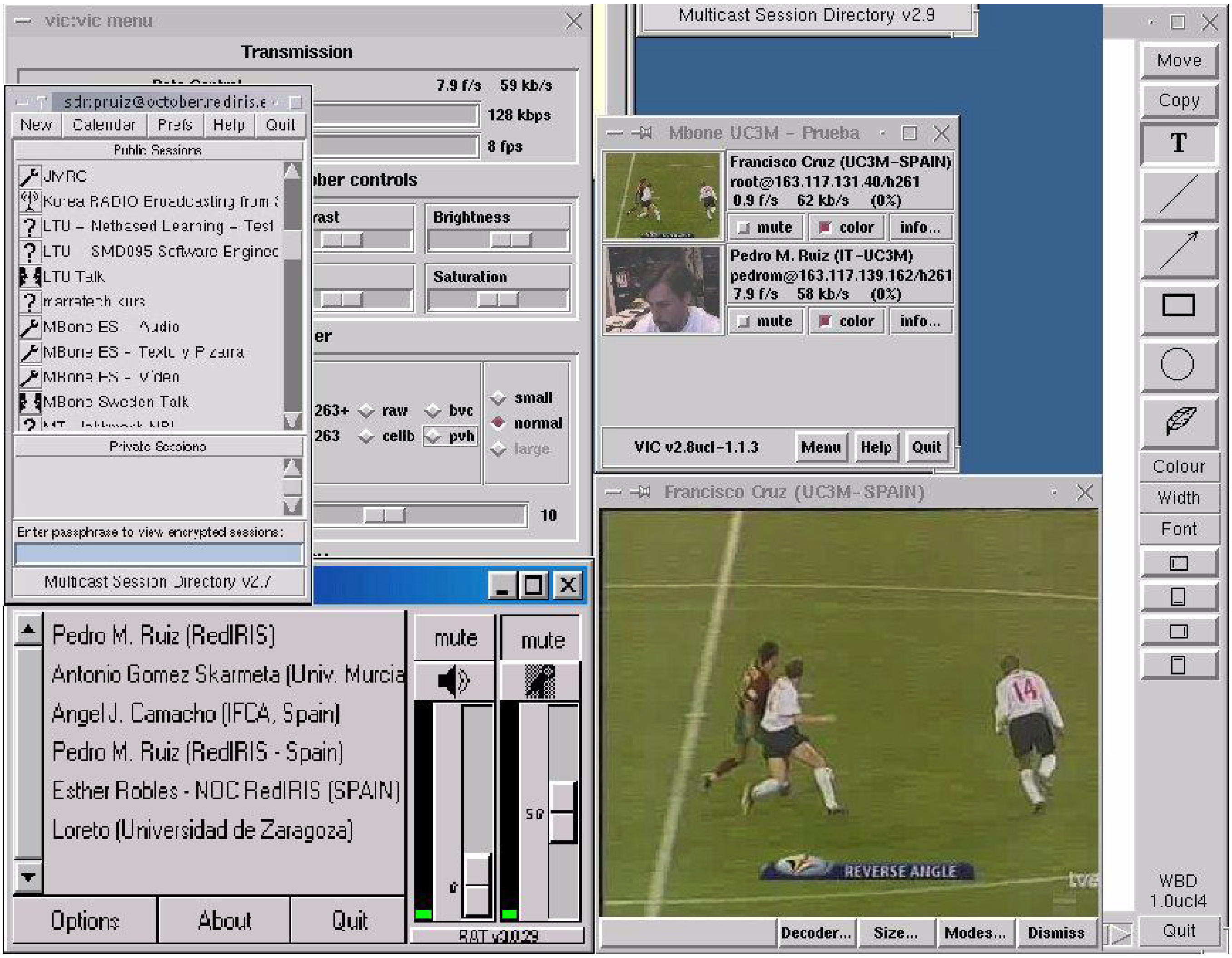


Figure 5: Some of the MBone tools

to Contents, page 1

4 Advanced services over IP Multicast

As mentioned in this article, video conferencing over IP Multicast initiatives led to an important number of applications and protocols. Many are used in other environments, mainly in VoIP architectures (i.e. RTP, SIP, SDP, codecs, etc).

However, IP multicast and its video conferencing tools have limitations that hindered the deployment of MBONE tools.

4.1 Multimedia services integration

Multimedia services over IP multicast present the following limitations:

- No integration with solutions based on some other technologies. It is desirable to have a solution to integrate IP multicast with H.320, H.323 and other VoIP solutions in general. SIP will be very useful as a glue element between all these technologies.
- No integration between the MBone tools. Currently, one different tool is used for every service (i.e. VIC for video, VAT for audio and so on). Although this model simplifies the design and development of the tools, it becomes an issue due to the lack of synchronization and user interface integration between tools. Nowadays, several proposals are coming up within the IETF to solve this problem. There are two basic models: mbus (Multicast Bus) and SCCP (*Simple Conference Control Protocol*).

4.2 Security and access control in multicast environments

In the same way that the video conferencing tools over IP Multicast present several limitations that are slowing its deployment, the limitations of the IP Multicast model are making ISPs to think twice before offering the IP Multicast service to their customers. The main problems of that model can be summed up as:

- Denial of Service (DoS) attacks.
- Policy of use, because there is not a defined method to control access to the network.
- Authentication.
- Address allocation.

To avoid these problems, several working groups at the IETF are defining new protocols or even updates to the current IP multicast model. Some such initiatives are:

- BGMP (*Border Gateway Multicast Protocol*). It is thought to be the successor of the currently used MBGP (Multiprotocol BGP). This protocol is very scalable and, if combined with other protocols for address allocation like MASC, is able to solve the address allocation problem.
- MSEC (*Multicast SECurity*). It is a new IETF working group that is responsible for studying and solving security concerns in IP multicast.
- SSM (*Source Specific Multicast*). This is a new multicast model based on the concept of *channel*: a pair of a source and a multicast group. The key concept in SSM is that routing decisions are taken based on channels instead of multicast groups.
- GLOP. This mechanism divides statically and according to the AS (autonomous system) number, the 233.0.0.0/24 range so that there won't be collisions between ASs when selecting an IP multicast group. Figure 6 shows how to know what groups belong to what Autonomous System. For example, in the case of RedIRIS, the AS number is 766, so the range 233.2.254.0/24 is available for being used within the RedIRIS AS without worrying about possible collisions.

0	8	16	24
233	16 bits AS		local bits

Figure 6: Structure of GLOP addressing

5 Conclusions

In this article, we have covered the most important topics related with IP Multicast technology and video conferecing over IP Multicast networks. We have showed how important protocols developed inside MBONE initiative have been later reused in the most outstanding VoIP proposals like H.323, SIP or MEGACO. Although multimedia over packet networks is a wide subject and has been vastly investigated and experimented, we can consider MBONE as an important testbed where basic technologies nowadays used in VoIP have been matured.

Although multicast video conferencing is not popular at this moment – most of present VoIP scenarios resemble the ones found in conventional telephone networks, and so, they are unicast –, IP multicast opens possibilities for new applications, improving the use of network resources. But, as mentioned before, much more development and research is needed to improve multimedia conferencing over IP Multicast.

References

[Deering 89]
S. Deering, Host Extensions for IP Multicasting, RFC 1112, 1989

[Estrin et al. 98]
D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, C. Liu, P. Sharma, L. Wei, Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification, RFC 2362, 1998

[Macedonia/Brutzman 94]
R. Macedonia, D. P. Brutzman. MBone Provides Video and Audio for Internet Across the Internet. IEEE Computer, Vol. 27, April 1994, pp. 30–36

[Meyer 98]
D. Meyer, Administratively Scoped IP Multicast, RFC 2365, 1998

[Moy 94]
J. Moy, Multicast Extensions to OSPF, RFC 1584, 1994

[Schulzrinne et al. 96]
H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, RTP: A Transport Protocol for Real-Time Applications, RFC 1889,1996

[Handley et al. 00]
M. Handley, C. Perkins, E. Whelan, Session Announcement Protocol, RFC 2974, 2000

[Handley/Jacobson 98]
M. Handley, V. Jacobson, SDP: Session Description Protocol, RFC 2327, 1998

[Handley et al. 99]
M. Handley, H. Schulzrinne, E. Schooler, J. Rosenberg, SIP: Session Initiation Protocol, RFC 2543, 1999

[Waitzman et al. 98]
D. Waitzman, C. Partridge, S. Deering, Distance Vector Multicast Routing Protocol, RFC 1075, 1998

to Contents, page 1

Implementing Voice over IP

André J. Hes and Ronald van Teeffelen

This paper describes the implementation of Voice over IP services on a wholly owned IP infrastructure. The benefits, minimum requirements and the sensitivity of IP data packages are outlined based on research and experience gained during large-scale pan-European wide Voice over IP implementations.

Keywords: VoIP, Gateway, Gatekeeper, packet loss, jitter, packet queuing, packet buffering

1 Introduction

VoIP allows customers to use the access service for voice in two ways. First, customers can use the PSTN breakout service that allows them to offload their H.323 VoIP traffic to the network, and have their telephone calls sent out onto the PSTN in the country closest to the dialled subscriber. This least-cost routing functionality reduces long-distance telephone costs for *off net* calling.

The second service for customers is the possibility to use the IP network for *on net* calls. Customer PBXs at various office locations throughout Europe are coupled via the IP network, enabling them to route calls between their offices across one network, via CPE gateways. In effect, the existing IP infrastructure is used for both voice and data traffic. It is also possible to run incoming calls on IP via gateways. This functionality is available for service numbers only (e.g. free-phone numbers).

VoIP, integrated with data traffic, creates a foundation for new possibilities that can significantly reduce cost for voice calls. This, in turn, opens up numerous possibilities for offering value-added services in this new integrated space. To take just one example, international call centres can use their resources more effectively using both language and skill-based routing, in conjunction with knowledge databases. In theory, integrated data and voice can fully handle the call with pre-recorded digitised messages by identifying your customer, and their potential question or problem. This provides a level of call intelligence, well in advance, of current call centre capabilities.

2 Minimum requirements

Customers are connected via the wholly owned IP infrastructure. The breakout possibility to PSTN is provided by the VoIP gateways. In general voice services can be provided over IP when:

- Customers have dedicated access to the network (minimum 256kbit/s connection required),
- Quality of Service (QoS) must be provided,
- Any firewalls or network address translation devices in the network path must be H.323 aware,
- Link layer QoS mechanisms should be in place for customer networks and networks that are not over-provisioned, as well as for the local tail,
- PSTN-Gateways will be located at the Points of Presence (PoPs).

Telephony systems supported by KPNQwest VoIP services must meet one of the following criteria:

- PBX with ISDN PRI or BRI interface
- PBX with E&M or analogous FXO interface
- Telephone with analogous FXS interface

3 Voice over Internet Protocol (VoIP); how it works

Figure 1 shows the general network layout of Voice over IP.

VoIP has two physical components:

- *PSTN Gateways:* Terminate calls from physical sources within the network to the PSTN. The gateways are located at PoPs in the IP backbone.
- *Customer equipment:* A VoIP gateway at the customer premises is used to connect a customer PSTN switch or PBX to the VoIP cloud. Usually, the customer also has a PSTN connection on his PBX and incoming calls are routed over this connection. The VoIP gateway can also be integrated in the Customer Premises Equipment (CPE) router. The gatekeepers provide registration, admission and status (RAS) functionality and Least Cost Routing services. Gatekeepers also control the VoIP switching logic. The centrally located

André J. Hes is Senior Project Manager at KPNQwest, currently working on the merging of four regional European call centres to one virtual centre. He has over 15 years experience in IT, as business consultant, Project and IT-Manager, Graduate in Business Information Technology, and Postgraduate EDP-Auditing at the Amsterdam Business School. Receiver of the KPNQwest Hero award for the Nokia-game project. Previous employers; Tandon, KLM, Atos-Origin, Philips. <andre.hes@KpnQwest.com>

Ronald van Teeffelen is currently working with KPNQwest as a Technical Consultant on VoIP within Technical Development. His main responsibility is the development of new VoIP services. He is also interested in IP VPN services. Van Teeffelen holds a Masters Degree in Information Engineering, has over five years of experience in networking. <ronald.vanteeffelen@kpnqwest.com>

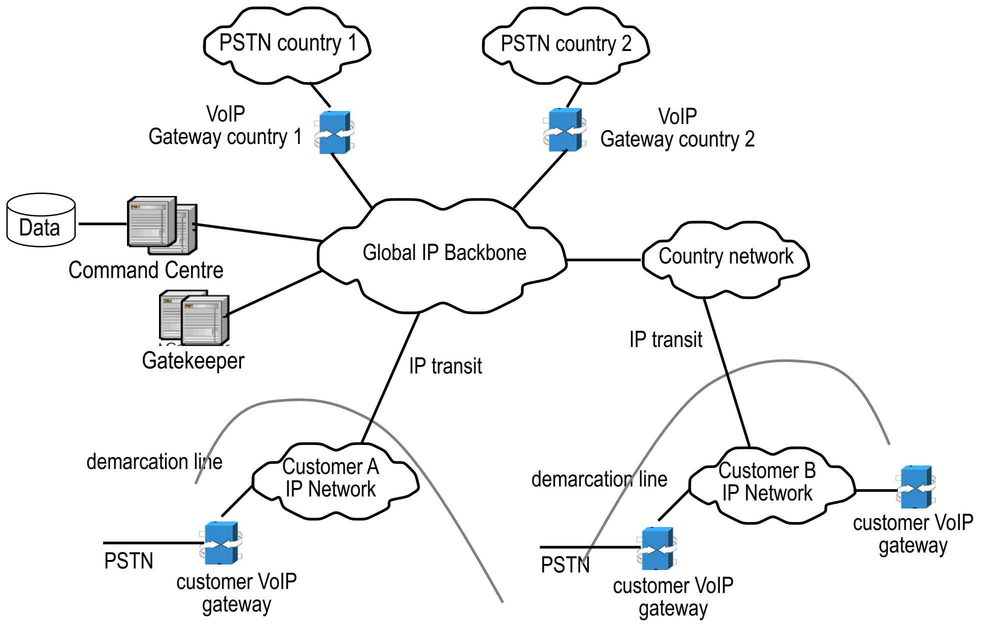


Fig. 1: General network layout of Voice over IP

Command Centres contain the routing tables. This set up allows a simple standard configuration of dial peers on the gateways and central management of the routing tables and Least Cost Routing service.

There are three basic types of traffic:

- *RAS messages between the VoIP gateways and the Gatekeeper*. These messages follow the standards H.323 and especially H.225 and H.245. To set up a voice connection, the initiator starts on an H.225 connection over TCP to the Gatekeeper at port number 1720. In this session a port number for the following H.245 connection is exchanged. The initiator then opens an H.245 connection to the gatekeeper over TCP (ephemeral port), in which ports for the actual voice traffic between two H.323 terminals (e.g. CPE Gateway and PSTN-Gateway) are exchanged (RTP traffic, ephemeral UDP ports). While the H.225 connection could be torn down after the H.245 ports have been exchanged, it will in practice stay up until the call is over.
  The Gatekeeper itself will also open connections to the terminating H.323 terminal in order to be able to negotiate the ports that should be used between the initiating and the terminating H.323 terminal.
- *Actual voice traffic*. The actual voice traffic flows over RTP between two H.323 terminals. For this reason, two RTP and one or two RTCP flow are established (over UDP). The ports to be used for these RTP flows have been exchanged in the preceding, (see above) H.245 connection.
- *Administrative traffic*, which includes administration (over telnet, ssh, http and https) of gateways and gatekeepers and logging of information like SNMP traps or call data records.

How the process works:

- When the CPE VoIP gateway receives a call, it contacts a gatekeeper. The gatekeeper locates the correct PSTN gateway to route the call to, and forwards the IP address of the PSTN gateway.
- The gateways then open up an H.323 call. In the event of a PSTN gateway failure, calls will still be routed to their destination.
- If one of the gateways fails, the gatekeeper automatically routes the call to a backup gateway.
- Calls will be routed to the gateway depending on the country code dialled. For example, calls to Germany will be routed to Frankfurt, and will leave the network from there as a local call.
- The gatekeeper uses authentication to prevent unauthorised usage. If the voice interface (PRI/BRI) to the gateway fails, the PBX will route the call over the PSTN.
- The backbone provides reliability and availability for calls in progress.

4 Voice data packet sensitivity

Unlike data traffic, voice traffic is intolerant of packet loss and delay, primarily because these conditions degrade the quality of voice transmission delivered to the end user. Delay must be constant for voice applications. The ITU-T G.114 recommendation specifies that for good voice quality, no more than 150 milliseconds of one-way, end-to-end delay should occur. KPNQwest, for example, guarantees 80 ms over its entire IP backbone.

IP Packets that are to be sent to a specific interface on a router are first put into a queue, from which they are sent. When queues fill up, this will add latency for packets that freshly enter the queue. This can happen if the sender is able to send data at a higher speed than the speed of the access line. In order to prevent this filling up of the queues having negative impact on the voice quality due to latency, alternative queuing strategies must be used, which prefer voice packets to other packets (Fig. 2).

In this scenario, a separate queue for voice packets exist. Voice and non voice packets are sent alternately on the wire, reducing latency for the voice packets.

5 Delay variation versus time: jitter

As VoIP is a packet switched service, it will show a variable delay in packet delivery time, known as jitter. Packets transmitted between Coder and Decoder through an IP network will experience different delays and will arrive at a non-uniform rate. The main causes for jitter are the buffering and

to Contents, page 1

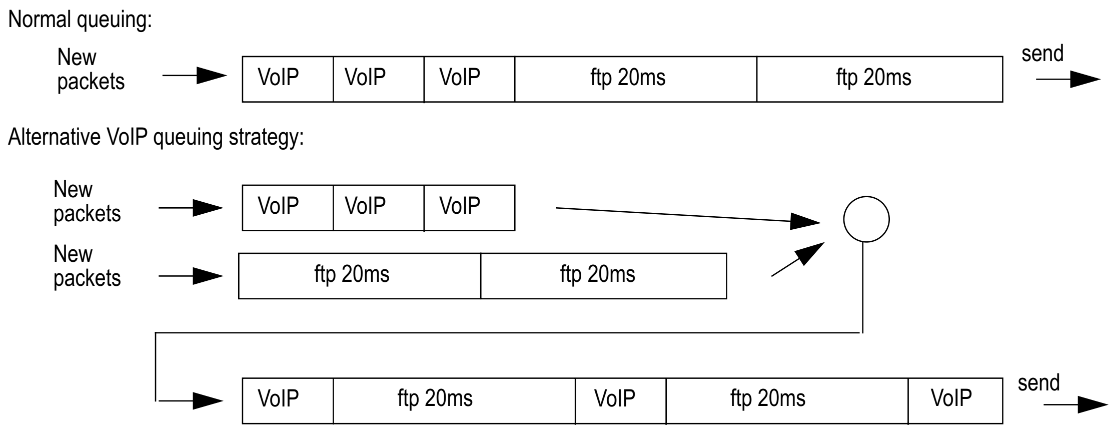


Fig. 2: Queuing strategies for voice packets

queuing processes in each of the IP 'hops'. In addition, the simultaneous transport of data through a relatively low-bandwidth local tail will cause jitter in the VoIP packets.

Since the Decoder must decode the IP packets into a constant voice stream, a buffering mechanism is needed. This buffer must be implemented at the receiving side in the gateway to absorb the fluctuation in packet arrival rate. The delay of an IP packet through the network is composed of a fix part characteristic of the propagation delays and the average queuing length, and a variable part characterising the jitter. The jitter buffer will hold arriving packets in memory for a certain time before sending them to the decoder. The larger this jitter buffer is, the more jitter on the packets can be absorbed. However, the delay introduced by the jitter buffer reflects in the one-way, end-to-end delay directly. In essence, a trade-off must be made between packet loss and introduced delay.

In practice, two types of jitter buffers can be implemented:

- *Dynamic jitter buffer:* terminals use heuristics to determine the size of the jitter buffer. Starting, for example with a small jitter buffer and progressively increase it until the packet drop is below a specified percentage.
- *Fixed jitter buffer:* The size of the jitter buffer can be set to a fixed value.

The effect of jitter will be most pronounced when voice and data is transported over a low band width link. A 1,500 bytes IP data packet transported over a 512 kbit/s link can cause a jitter in VoIP packets of approximately 1500*8/512k = 23 ms.

6 Other QoS Mechanisms

As previously mentioned, voice traffic is sensitive to packet loss and delays. This could be the result of an overfull connection on the local tail, the access line between the nearest Point of Presence (POP) and the Customer Premises Equipment (CPE router). To prevent this, quality of service (QoS) mechanisms need to be put in place.

In order to achieve an optimum voice quality, the POP and CPE router will have to be capable of treating voice traffic in a different manner to the rest of the network data. In addition, the minimum requirement for a local tail of 256 kbit/s for a VoIP customer, and to ensure that delay sensitive voice traffic is treated with priority, some of the following queuing mechanisms need to be used:

- IP Real Time Protocol (RTP) compression will be employed if the local tail bandwidth is below 2 Mbit/s (E_1 speed); the delay overhead for compression quashes any gains when link speeds are over 2 Mbit/s and can only be implemented on PPP, ISDN, HDLC and frame relay links.
- IP RTP prioritisation will be used to provide a strict priority queuing scheme for voice traffic.
- Class-Based Weighted Fair Queuing (CBWFQ).

One of the intrinsic values of WFQ in the data-only world is its fairness. WFQ is designed to fairly share available resources between bursty traffic types. This particular aspect of WFQ, that is so beneficial to data, is what renders it inadequate for packetised voice traffic. Although voice traffic can be assigned an IP Precedence of 5 to give it a weight that grants it greater priority than other flows, if a large number of competing flows exist, voice traffic may not be allocated sufficient bandwidth to maintain the required kind of quality. The quality of controlled constancy of delay and packet loss is required all the time, regardless of the priority assigned to the packet. Therefore, to ensure that VoIP traffic always gets the bandwidth it requires, IP RTP Priority (and other features) can be used for voice traffic, and CBWFQ for all other traffic.

7 Other issues

Echo cancellation

Feedback chains in analogous equipment like the telephone handset can impose a certain echo that travels back up the line to the original speaker. The VoIP gateways can do a certain amount of echo cancellation by "remembering" the last few milliseconds of speech and subtracting these from the signal coming back to the gateway. If the local loop (the distance from the analogous interface to the connected equipment producing the echo) is longer, the value of time to be remembered should be extended.

Codec	Voice Bandwidth	Transport Bandwidth	Default Packet size	Packetisation delayMs	Coding Delay Ms	Complexity (according to Cisco)	Quality	MOS value	KPNQwest VoIP
G.711	64 kBit/s	80 kbit/s (with RTP header compression about 65 kbit/s)	160	20	0.375	Low	Normal	4.1	✔
G.729	8 kbit/s	12 kbit/s (RTP comp.) 24 kbit/s (without rtp comp.)	20	20	35	Medium	Normal	3.92	✔
G.726	16 kbit/s	32 kbit/s (with RTP header compression about 17 kbit/s)	40	20	0.375	Medium	Analog		✔
G.726	24 kbit/s	40 kbit/s (with RTP header compression about 25 kbit/s)	60	20	0.375	Medium	Normal		✔
G.726	32 kbit/s	48 kbit/s (with RTP header compression about 33 kbit/s)	80	15	0.375	Medium	Normal		✔
Fax, T.38						Medium	N/A	N/A	✔

If a longer time is configured for this feature, it will take the echo canceller longer to converge; in this case, the user might hear a slight echo when the connection is initially set up. If the configured value for this command is too short, the user might hear some echo for the duration of the call because the echo canceller is not cancelling the longer delay echoes. The voice ports can be configured for input gain, output attenuation and echo-cancellation.

Silence detection

In order to reduce network bandwidth, the VoIP gateway can detect pauses in speaking and not transmit data for the duration of these pauses. This results not only in reduced bandwidth, but also in a sort of 'total silence' for the receiver that is uncomfortable for most humans. The VoIP gateway is also capable of producing some "comfort noise". The generation of comfort noise should be turned on, if silence detection is turned on. In practice, silence detection should be turned off, unless bandwidth is restricted. By default, silence detection and comfort noise is enabled, so VAD must be disabled in the VoIP dial peer.

Clocking

PSTN equipment needs to be synchronised in order to be able to achieve a common sense on the understanding of time slots. For this purpose, a PBX gets clock information from the public switching equipment.

Signalling

For signalling within the H.323 cloud, all H.323v2 signalling standards as Q.931, H.323, H.245, H.225 and RTCP are supported. For signalling towards the PSTN side, ISDN BRI and PRI (EDSS-1) are supported for digital connections. For analogous connections, standards supported are E/M, FXS (only low-end version of product) and FXO.

General Voice Codecs

G.711 with A-law encoding is the default codec; customers can use G729 or G726 as well.

Voice gateways support different numbers of simultaneous calls depending on the codec complexity. Only for medium or low complexity codecs, the full number (equals to the number of access channels) of simultaneous calls are supported. For high complexity codecs, the number is lower.

8 Conclusion

Besides the minimum requirements stated above, due to the sensitivity and intolerance of VoIP data, maintaining a VoIP network and benefiting from cost-savings and future applications through combining voice and data over one network, requires a well-equipped and skilled organisation to guarantee optimal voice transmission.

to Contents, page 1

Voice over IP Virtual Private Networking

Olivier Hersent

This article describes the various aspects of the integration of Voice over IP (VoIP) into a Virtual Private Network. Key factors of deploying a VoIP Virtual Private Network service are pointed out, showing the service provider and the customer point of view. Value-added services such as off-net to on-net connectivity and unified messaging are outlined. In addition, architectural strategies of dialling plans and group prefixes are summarised. Quality of Service is also addressed.

Keywords: Voice over IP, IP VPN, PSTN, iPBX, number expansion strategy, Quality of Service

1 Introduction

Organisations around the world are constantly seeking to reduce communications costs and increase organisational productivity. Significant improvements in just these areas can be made today by using VoIP. When it comes to value-added services, especially in a distributed environment, VoIP is much more efficient than circuit-switched networks. It allows usage of a network that not only converges voice, data and video, but also facilitates the introduction of new applications.

The most crucial feature of a VoIP Virtual Private Network (VPN) is the ability to connect many corporate sites to a single network while preserving a virtual isolation of each group that communicates on the shared infrastructure. Each group can use optimised private dialling plans, relying on the VoIP VPN to manage the communications among themselves as well as public networks and groups.

The VoIP VPN is an application level overlay on top of any IP connectivity technology, and can be deployed with or without a data level VPN using technologies such as IPSec (IP Level Security Standard).

2 Deploying a VoIP VPN service

When selecting value added services, a change on the customer premises should be stringently avoided. Replacing the leased lines between Private Branch Exchanges (PBX's) – or the equivalent Intelligent Network Closed User Group service – with a VoIP VPN is an ideal example of a service that requires no change on the customer's side.

Technically, the VoIP VPN service does not require significant changes in the VoIP network architecture. It is an exact replacement of existing Public Swtiched Telephone Network (PSTN) technologies. The initial investment is therefore limited to one CCS (Call Control Server) Softswitch capable of VoIP VPN operation, and the installation of dedicated or shared VoIP gateways connected to corporate PBX's.

The VoIP VPN service is also an excellent means of introducing new VoIP value-added services in the future. For example, remote workers can be included on the corporate VoIP VPN. Advanced number routing and virtual call center services can be offered as well.

2.1 The VoIP service provider

Most of the investment that VoIP service providers need to make is used to purchase network infrastructure (such as VoIP gateways) and to install this infrastructure in multiple points of presence. This includes a basic call control infrastructure in order to route phone-to-phone calls. This infrastructure may also provide simple prepaid telephony and PC-to-phone services. Initial services usually use about 20–40% of the deployed capacity.

The opening up of the VoIP network to value added services is an opportunity to use some of the spare capacity of the deployed network for applications that are less sensitive to price. The VPN feature allows the core VoIP network to behave as a hub aggregating many different telephony services on the same backbone: corporate PBX's, corporate or hosted Internet Protocol Private Branch Exchanges (IPBX's), residential gateways and IP phones can all be connected to the VoIP VPN core. This enables VoIP Service providers to offer services to the whole range of end-user devices, from blackphones on PBX's to IP Phones on IPBX's.

Most corporations are permanently connected by data connectivity service providers for their e-mail, intranet, extranet and Internet services. Corporations with multiple sites already have a need for top quality connectivity services. These same corporate customers are also those who probably already have

Olivier Hersent, NetCentrex Founder and Chief Executive Officer, started NetCentrex in 1998. Previously, he was the lead VoIP architect for France Telecom R&D laboratories (CNET) where he focused on IP security, Multicast and QoS. Hersent is an active player of the VoIP industry through his participation in standards organizations such as ETSI TIPHON and technical conferences. He a published on the subject of VoIP, IP Telephony, Packet-based multimedia communication systems (Addison Wesley Longman). Hersent graduated from *Ecole Polytechnique* in Paris and completed a postgraduate telecom engineering program at *Telecom Paris*. <oh@netcentrex.net>

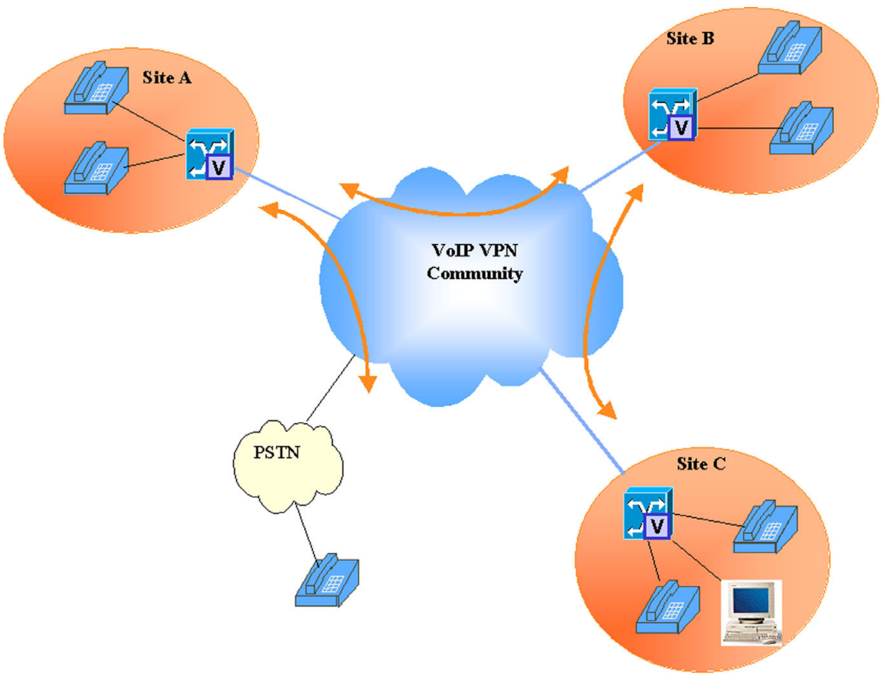


Fig. 1: VoIP VPN architecture

a voice VPN in place, albeit using traditional circuit switched technology.

2.2 The VoIP customer

By converging both voice and data services on the data network, all internal, site-to-site calls use only IP resources and are therefore included in the current monthly fee already paid for the data network. The previous architecture of leased telephony lines disappears. In addition, off-net calls are no longer routed through the corporate site that was closest to the destination, as in the previous circuit-switched architecture. With the VoIP VPN architecture, all off-net calls are routed through the VoIP gateways of the service providers, and use the routes and termination gateways chosen by the service provider. A change to the internal telephony infrastructure already in place is not required.

Opening the voice network to IP services also brings many new network-based services such as video conferencing, data sharing, virtual call centers, click-to-call back and click-to-call.

3 VoIP VPN Services

3.1 On-net to on-net calls

The core feature of the VoIP VPN service is to allow corporation sites to call each other as well as home workers over the voice intranet. The IP network carries all telephony traffic among corporate VPN sites and home workers. The architecture is also much more flexible when an office relocates, mainly because the same telephone numbers can be retained. VoIP also facilitates the integration of various brands of PBX's.

The VoIP VPN architecture also allows the definition of a simplified dial plan that can span several different companies with a voice extranet.

3.2 On-net to off-net

The core VoIP VPN architecture should include a set of least cost routing features allowing selection of the best terminations according to load conditions and prices. The VoIP VPN uses the home DSL or cable connection to extend the VoIP VPN reach into the home office. Single-step dialling is possible, and the user experience is the same regardless of the size of the office. There are no local call costs, and each user benefits from the discount plans negotiated for the entire VPN.

3.3 Off-net to on-net

Basic off-net to on-net connectivity allows endpoints in the voice VPN to be called from the PSTN using direct extensions. A VoIP VPN provides additional services such as Virtual Local Extensions and Voice Remote Access Interactive Voice Response (IVR).

For multinational companies, travelling workers frequently need to call their home office. By allocating virtual local phone numbers for some VoIP extensions, the PSTN no longer has to route calls to the home office in several steps over the PSTN. For example, it allows people travelling in Spain to call a local Spanish number rather than calling a long-distance number to their home office in another country.

The virtual attendant service, Voice Remote Access IVR, allows travelling employees to dial a specific number that asks them which internal VPN number they wish to call. With an optional identification procedure, they can also be allowed to make off-net calls at no charge or on a prepaid account.

3.4 Additional services

A "Subscriber Policy Engine" service can be very beneficial and can help reduce the difficulty of reaching people on the move by providing "follow-me" type of services. Roaming users simply need to be allocated numbers in a recognisable range. This service can be offered by the service provider as a specific class of service for roaming users, without requiring any specific installation within the core VoIP VPN architecture.

"Unified Messaging" allows users to retrieve and manage their voice mail as part of routine e-mail checks. This saves time and enhances the responsiveness of the entire organisation. A Unified Messaging system can easily be added as part of the core VPN architecture, i.e., the offering of the service provider, and should not require additional on-site hardware.

to Contents, page 1

4 Architecture

4.1 Overlapping dialling plans, group prefixes

With the number expansion strategy, the CCS Softswitch learns the physical location of an IP phone through the registration process. In the H.323 standard, the phone sends a registration request message to the gatekeeper with its alias (name or telephone number) and current IP address.

For example, if both groups *A* and *B* have a '1010' extension, the CCS Softswitch needs to know that whenever a member of group *A* dials '1010', the call needs to be routed to the '1010' phone within group *A*, not group *B*. The correct strategy for implementing a VoIP VPN is to make the phone alias unique by adding a group prefix to each short telephone number.

If group *A* is assigned prefix 111 and group *B* is assigned prefix 222, then:

- the '1010' IP phone within group *A* registers as 1111010
- the '1010' IP phone within group *B* registers as 2221010

When dialling out from a phone in a VPN group, the H.323 phone will use its registered alias as the Calling Party Number information element of the H.225 Setup message used to establish the call. By looking at the group prefix of the calling party number, any call from group *A* can be recognised.

Similarly, the corporate gateways connected to PBX's must add the group prefix to the calling party number for any call routed to the VPN network. This requires from the CCS Softswitch implementing the VPN:

- Simultaneous Source and Destination-based routing (ability to also route a call based on the calling party number)
- Calling Party Number and Called Party number expansion and truncation.

A variation in this configuration can be, for example, on networks with no IP phones (only gateways), it is possible to have a gateway add the group prefix to the Called Party Number instead of to the Calling Party number. This allows the implementation of a VPN network with a gatekeeper that does not have source-based routing capabilities.

Another strategy is to virtually isolate a VPN by using the list of H.323 ID's of the gateways participating in the VPN. The gatekeeper is provided with a list of H.323 ID's for each gateway. Therefore, whenever a call arrives from that gateway, it can decide to which VPN the call belongs and route it accordingly.

4.2 Communications with the public network

Most VoIP VPNs need to be reachable from the PSTN as well as be able to dial out to the PSTN. Dialling out to the PSTN is easy and only requires an escape access code that prevents mixing short VPN numbers and short PSTN numbers. In Europe, this is usually the digit '0'. The only difficulty is in properly choosing the egress gateway and in rewriting the Calling and Called Party numbers accordingly.

- In the Calling Party number, the VPN prefixes must be replaced by the PSTN prefix.
- In the Called Party number, the escape access digit must be removed and the number needs to be rewritten according to the local PSTN dial plan applicable in the country of the selected egress gateway.
- If a sequence of Direct Inward Dialling has been allocated to the VPN, then it is also possible to dial-in. When the gatekeeper receives a SETUP message from a gateway, it needs to rewrite the Calling and Called Party numbers.
- The Calling Party number needs to be prefixed with the appropriate escape code.
- The Called Party number needs to be rewritten using the VPN group and site prefixes.

The PSTN gateways are now shared. Some VPN groups may require a specific number of ports to be allocated for their own use.

4.3 Scalability, Upgradability, Availability

The technologies that are used on the core VPN platform are determined by the need of service providers to reach a critical traffic volume and negotiate better termination tariffs. The platform should support hundreds of calls per second while maintaining a very low call set-up time (in the 300ms range at most) at peak-loads. This precludes the use of relational databases in the softswitch call processing. Relational databases can be used only for the provisioning architecture, but the softswitch itself should use only in-memory, real-time telecom-grade databases. User profiles should be retrieved from read-optimised servers.

Very specific route indexing techniques, which guarantee a call set-up time largely independent of the number of routes, should be used. The provisioning framework should allow complete automation of operations for adding and removing sites. This makes it necessary to offer a command line interface, in addition to graphical user interfaces, since most provisioning database frameworks use a shell interface with third party systems. The softswitch system has to be hot-upgradable,

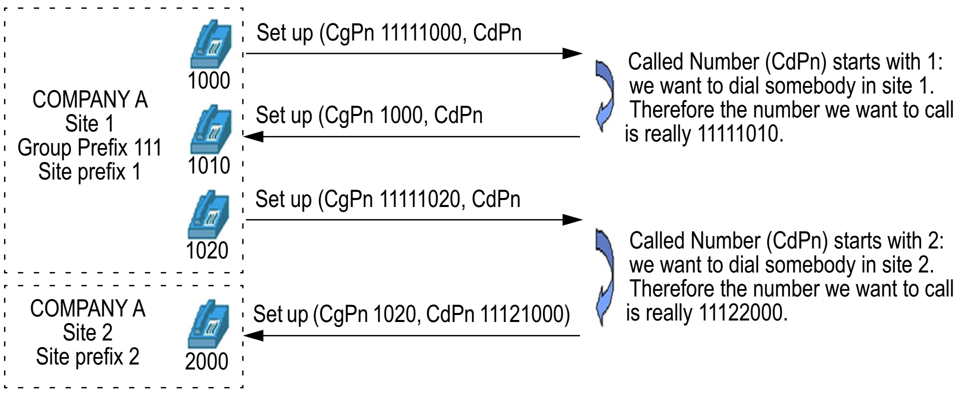


Fig. 2: A call example

both in terms of capacity and in terms of software release. Services should never be interrupted.

The softswitch system should be able to resist any single component failure, including software switches. A monitoring system should be able to detect and report hardware, operating system and application failures in almost real-time.

4.4 QoS – Quality Of Service

If transit delay in the IP network and packet loss can be controlled, the voice quality of VoIP can be comparable to that of PSTN.

In the corporation itself, on-site VoIP equipment and regular data hosts should use two separate levels of QoS. In a Local Area Network (LAN) architecture, the disruptions to a host by traffic originating from another host can be minimised. Using switches or Virtual LANs already provides a sufficient level of isolation. Under bandwidth constraints, voice packets should be prioritised using a priority scheme for all User Datagram Protocol (UDP) packets, or only for Real-Time Protocol (RTP) packets. For the leased line between the corporation and the core IP network, since it is usually under stringent bandwidth constraints, it is vital to use priority schemes for voice packets.

At the network peering points, the best situation would be one in which all used the same IP network providers. If not, dedicated symmetrical bandwidth should be provided in order to support the peering agreement among the networks. Public peering nodes and satellite links should be avoided.

Monitoring the physical integrity of all VoIP VPN components is essential in order to ensure a high service availability. A proactive monitoring approach, which continuously verifies the availability of the service at the application level, and can provide near real-time reports to administrators, is preferable.

5 Conclusion

Using VoIP in place of circuit-switched networks can create significant savings. In addition, the VPN feature allows the core VoIP network to behave as a hub aggregating many different telephony services on the same backbone to offer services to the whole range of end-user devices. Quality of service is comparable to PSTN, with additional services to come. Opening a telephony network to IP services affords the opportunity for many additional, new, network-based services.

VoIP in Public Networks: Issues, Challenges and Approaches

Francisco González Vidal

The present article introduces the basic problems found for conveying voice in a public environment, providing the quality of service the subscribers are currently used to. We schematically present the current methods for providing an estimation of the voice quality, as well as the performance of the current coding schemes and their interworking. A discussion about public network architectures based on VoIP follows, as an evolution scenario from the current Public Switching Network, Time Division Multiplex based, to the future Next Generation Network, Packet based. The network elements of the new paradigm are introduced: Softswitch, Trunk Gateway, Access Gateway. Finally, a brief description of a current project on Distributed Access Gateways for supporting the above described evolution scenarios is presented.

Keywords: VoIP, QoS, Access Gateway, Next Generation Network

1 Towards the Next Generation Network

There is an industry consensus that the telecommunication network of the future will have the following main traits:

- Network intelligence will migrate to the periphery, either to the user terminals or to servers where high level applications are located.
- There will be only one multiservice telecommunication network supporting all the current and future telecommunication services.
- Data applications will be dominant with regard to bandwidth consumption.
- As for the current state of the technology, IP is considered as the network protocol of choice, the federating protocol.

Although the debate is open on some of these points, e.g. when video services were widely deployed would still data applications be so dominant in bandwidth? what kind of video services will be deployed? But the shift of the present voice centric Time Domain Multiplexed network towards a packet paradigm seems to be clear. What it is still not clear is the pace and the shape this major change will take. In this article we address the challenges and issues that have to be confronted in offering a voice service, more specifically, a voice service up to the standards of current telephonic service in this new packet oriented scenario.

The paper is organized as follows. First, we present the quality issues that arise in conveying voice over packets, specifically on IP that seems to be the technology of choice. Then, we address the proposed architecture for the Next Generation Network (NGN) from the point of view of the voice service. Finally, we describe a current project we are involved in Alcatel for providing a Distributed Voice over IP Gateway on top of an Integrated Multiservice Access Platform (IMAP). This approach is aimed at easing the migration from the current Time Division Multiplexing (TDM) network paradigm to the NGN in a way that keeps the vast investment that operators have to do in the access environment.

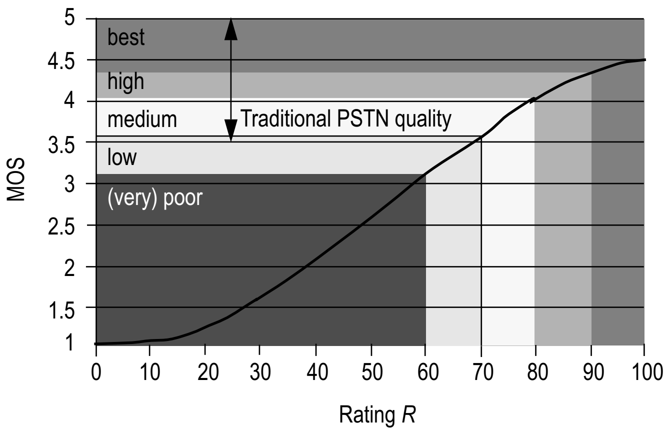


Figure 1: Mean Opinion Source (MOS) and Rating factor (R) are used to measure subjective voice quality

Francisco González Vidal is Telecommunication Engineer, and has a Ph.D., in the *Universidad Politécnica de Madrid*. He is Associate Professor in the Information and Telecommunication Technology Department ETSI Telecommunication in Madrid. In 1973, he joined Standard Eléctrica S.A., now Alcatel Spain, where he has developed his expertise in the Switching division (S12 and Broadband ATM). Presently, he is worldwide technical Director Access Systems of Alcatel.

to Contents, page 1

2 Voice Quality issues in an IP network

Voice quality is by nature a subjective matter that is measured by a set of currently standardised subjective tests. At present, two main indicators are used to measure voice quality:

- Mean Opinion Score (MOS) in a scale of 1 to 5
- Rating factor (R) in a scale of 0 to 100

Figure 1 shows how these two indicators relate to each other. We can appreciate that voice quality in traditional Public Switched Telephony rates from medium to best in performance. The subjective quality of a telephone call as a function of the network parameters is predicted with a model, called the E-model [ITU-T G.107]. The E-model is expressed as:

Origin	Standard	Type	codec bit rate (kb/s)	voice frame (ms)	look ahead (ms)	algorithmic delay (ms)	Ie	Intrinsic quality
ITU-T	G.711	PCM	64	0.125	0	0.125	0	94.3
	G.726 G.727	ADPCM	16				50	44.3
			24				25	69.3
			32				7	87.3
			40	0.125	0	0.125	2	92.3
	G-728	LD-CELP	12.8	0.625	0	0.625	20	74.3
			16				7	87.3
	G.729(A)	CS-ACELP	8	10	5	15	10	84.3
	G.723.1	ACELP	5.3	30	7.5	37.5	19	75.3
		MP-MLQ	6.3				15	79.3
ETSI	GSM-FR	RPE-LTP	13	20	0	20	20	74.3
	GSM-HR	VSELP	5.6	20	0	20	23	71.3
	GSM-EFR	ACELP	12.2	20	0	20	5	89.3

Table 1: Intrinsic characteristics of standardized codex

$$R = R_0 - I_s - I_d - I_e + A$$

Where:

- R_0 groups the effect of noise, either background or circuit noise
- I_s includes impairments simultaneous to the voice signal: due to quantization, too loud a connection, too loud side tone
- I_d encompasses delayed impairments, included those caused by talker and listener echo or loss of interactivity
- I_e covers the impairments caused by the use of special equipment; for example, each low bit rate codec has an associated impairment value. This impairment value can also be used to take into account the influence of packet loss.

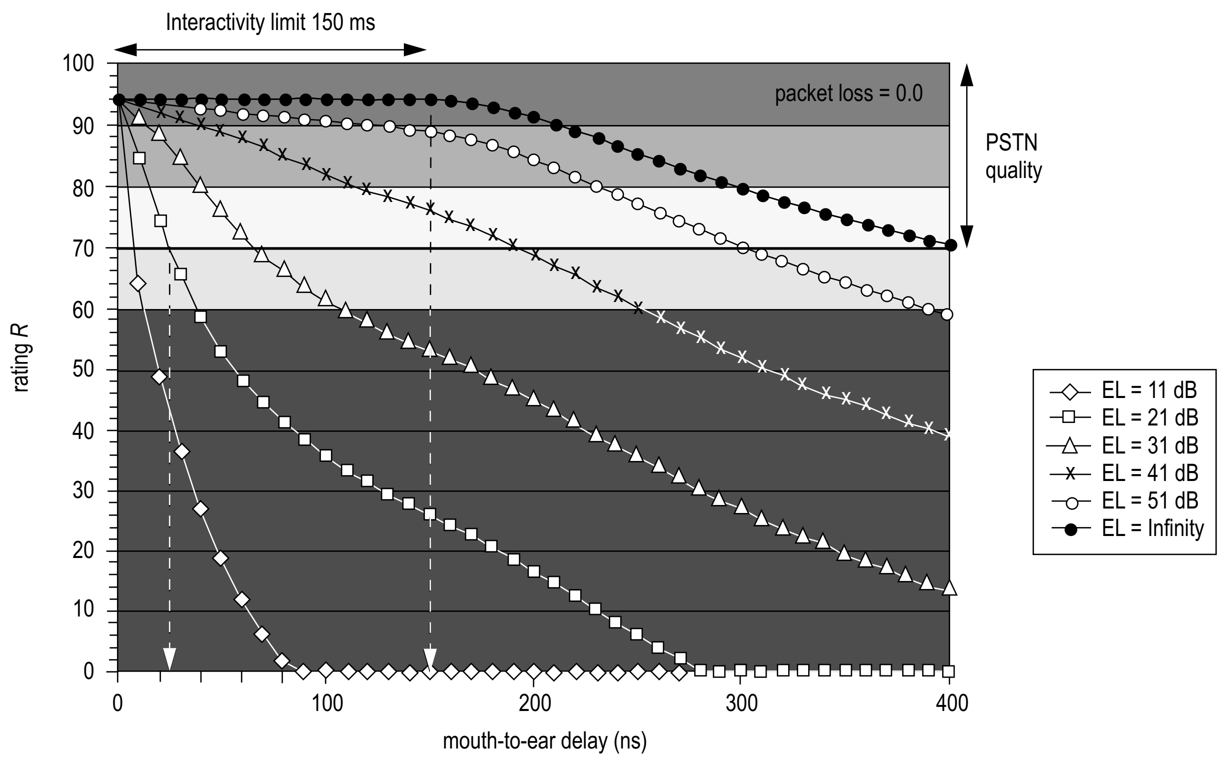


Figure 2: Effect of echo attenuation (EL = Echo Loss) on the rating R assuming no packet loss

to Contents, page 1

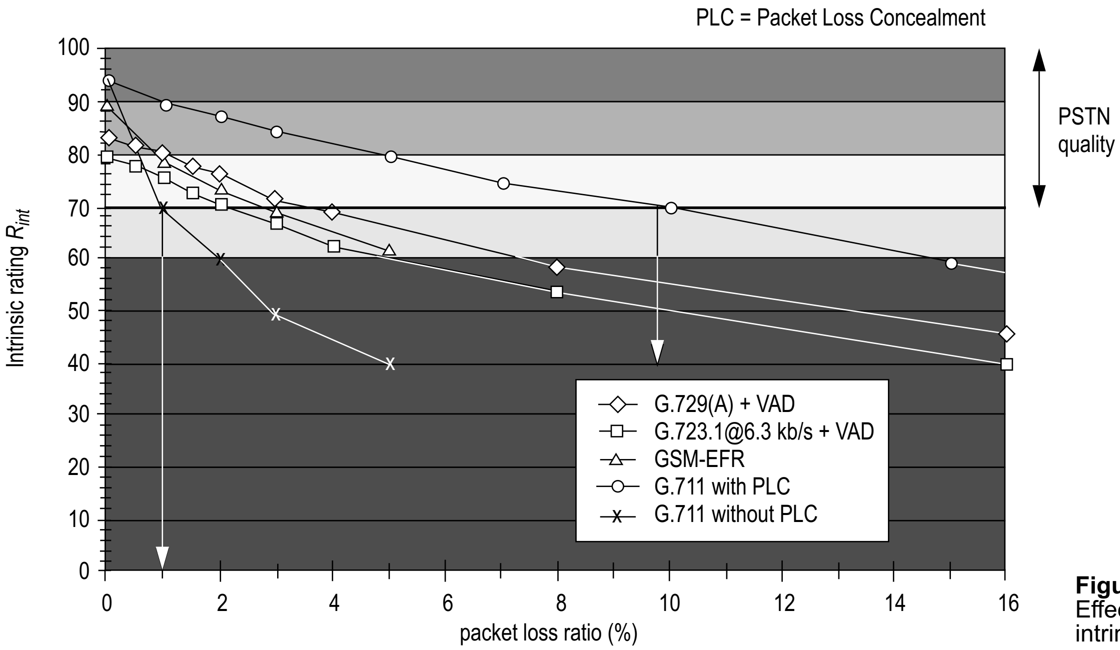


Figure 3: Effect of packet loss on the intrinsic performance

- Term A is the expectation factor. It expresses the decrease in the rating R a user is willing to tolerate in favour of the "access advantage" some systems have over wire-bound telephony. As an example, the expectation factor for mobile telephony is 10.

Two impairments mainly hamper the subjective VoIP quality:

- The impairment I_e associated with the use of "special equipment", like coding or transcoding systems, or distortion, like the one produced by packet losses.
- The impairment I_d associated with the mouth-to-ear (M2E) delay, produced, for example, by the loss of interactivity or the echo (talker and listener echo).

With regard to I_e, Table 1 shows the intrinsic characteristics of standardized codecs. This table shows that the intrinsic quality of most coding schemes is between medium to best performance (exceptions are ADPCM 16 and 24 kbps). Simulations performed in [De Vleeschauwer et al. 00] and [De Vleeschauwer et al. 99] show that transcoding is something that must be avoided as much as possible.

I_d is an important factor to take into consideration in packet based networks, as delay will be always higher than in conventional public telephony networks. The delay introduced when using a packet networks to carry voice traffic is made of:

- encoding-decoding delay (table 1), that is, the time needed to code and decode a set of voice samples using a codec algorithm,
- packetization delay, that is, the time needed to build a voice packet. For example, in G.723.1, each voice packet carries 30 ms of digitized voice (table 1),
- propagation delay, that is the time needed by electrical or optical signals to propagate (most important in the case of satellite links),
- queuing and transmission delay, that is, the time taken to transmit packets through the network, and
- dejittering delay, that is the delay introduced by the use of buffers at the destination to absorb delay variations in packet networks.

We have seen that I_d covers the impairments caused by talker and listener echo, and the loss of interactivity. The first ones can be treated by introducing echo control.

In Figure 2, we show the effect of echo attenuation (EL = echo loss) on the rating R for G.711 coding, assuming 0% packet loss. Note that, from 150 ms on, the rating decreases, due to the perceived loss of interactivity. Values above 400ms are unacceptable.

Finally, in Figure 3, we can see the effect of packet loss on the codec's intrinsic performance, i.e. without accounting for additional impairments. The beneficial effect of packet loss concealment, i.e. a strategy to replace the lost packets is also shown. A more detailed discussion of the combined effects of impairments can be found in [De Vleeschauwer et al. 99].

Finally, we can conclude that offering interactive VoIP with PSTN quality is perfectly feasible for a set of standardized codecs provided that:

- *perfect echo control* is performed as close as possible to the source of the echo (a hybrid in the PSTN or a PC terminal with low coupling loss),
- *transcoding steps are avoided* as much as possible,
- *IP network resources* for voice are *decently managed* (so that packet loss is low or negligible)
- *Mouth-to-ear delays* are *bounded*, that is:
  – voice packets get head-of-line over data packets in both access and backbone networks,
  – long data packets are segmented or interrupted as soon as voice packets arrive (e.g. in low speed access networks)

– dejittering and packetization delays are properly engineered.

3 Next Generation Network Reference Model

In this section we concentrate on the network elements needed to provide the telephony service. We leave aside elements pertinent to other services like data and video, although some of the elements can play a role in multimedia scenarios that include voice.

Figure 4 represents the Next Generation Network (NGN) reference architecture for providing voice services. Note that we have included in the picture the current Public Switched Telephony Network (PSTN), as well as the current telephone sets, as they are going to remain in service for any sensibly foreseeable future (keep in mind that PSTN world-wide has more than one billion main lines today).

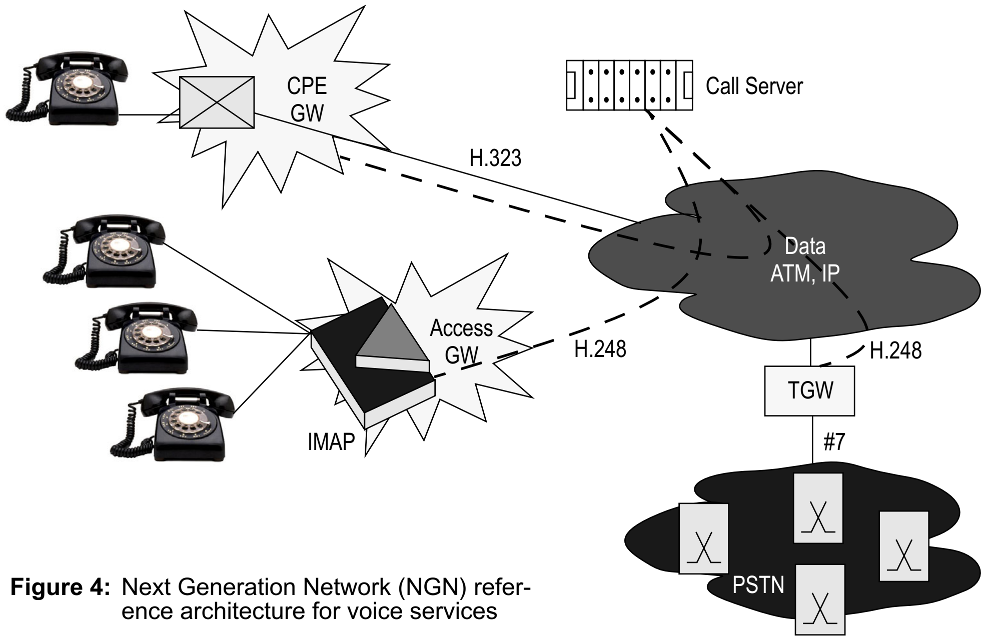


Figure 4: Next Generation Network (NGN) reference architecture for voice services

In Figure 4 we can identify the main components of this new scenario, namely the Gateways and the Call Server:

- Trunk Gateway (TGW): it performs at the bearer plane the necessary transcoding functions from the coding used in the PSTN network (normally PCM G.711) to whatever coding scheme we have decided to use in the packet network. In addition, at the control plane it has to perform the signalling gateway function between the legacy PSTN signalling system, (nowadays normally based on ITU-T Signalling System N. 7) and the signalling system chosen for the NGN, H.248 in the figure.

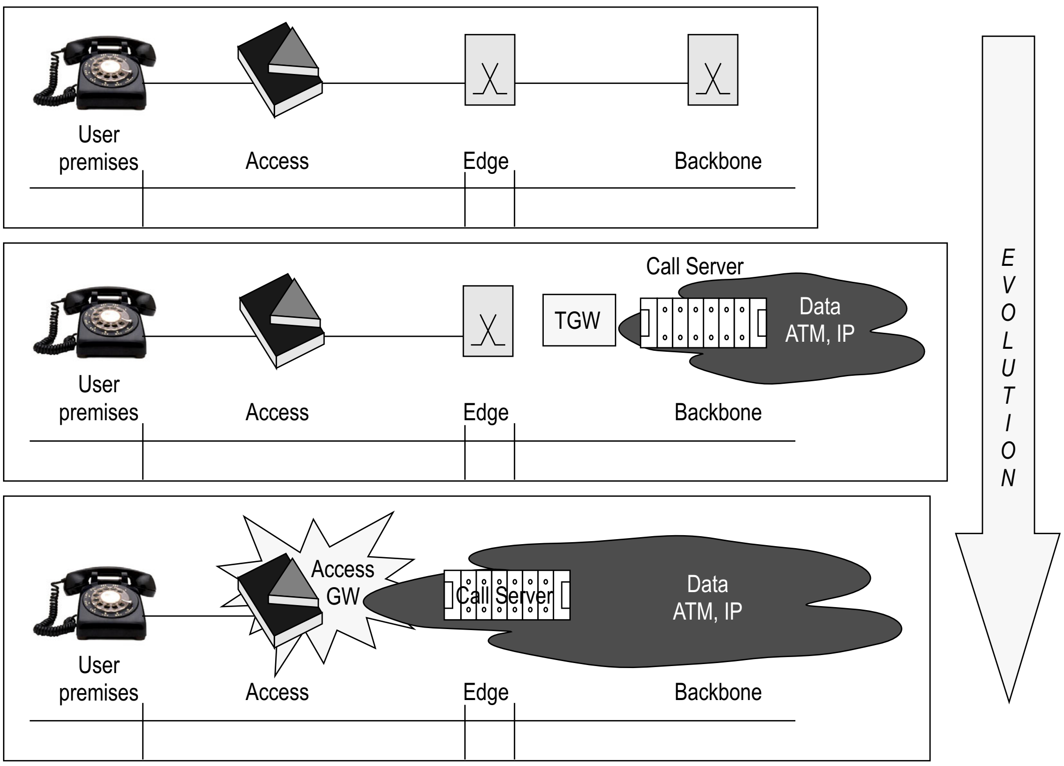


Figure 5:
Migration from the current network towards NGN

- Customer Premises Equipment Gateway (CPE GW): this is the VoIP application residing on the customers premises that will be used in the final scenario when all voice sets have become "IP phones". In a near future VoIP gateways will be embedded either in IP-PABX or in Integrated Access Devices, both for business users (of a significant size we should add). For the time being, VoIP Gateways for residential users are not economically reasonable.
- Access Gateways: these elements concentrate users having standard telephonic gear (analogue or digital phone sets, fax machines, key systems, analogue or digital PABX). It performs the corresponding gateway function both in the bearer plane, as we saw in the Trunk Gateway, as well as in the control plane, where the signalling conversion is done between subscriber signalling system (either analogue or ISDN) and the NGN signalling scheme, H.248 in this example.

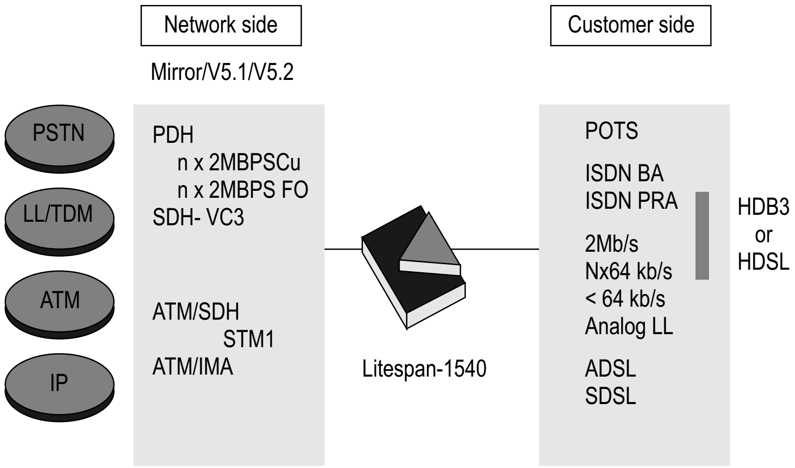


Figure 6: Integrated Multiservice Platform (IMAP)

To make this happen, a sensible migration path from the current network towards the NGN has to be planned. Current industry thinking favours an evolution starting in the backbone and progressing towards the user. Figure 5 provides a coarse view of this strategy.

Finally, we would like to add some considerations on the evolution scenario presented in Figure 5. In the case of new operators without a large TDM installed base, some intermediate steps in the evolution can be skipped. Of course, this will depend on the starting point of the new operator, as almost none of them start from scratch (typically all of them already have a TDM based backbone network).

If they come from the long distance interconnection business, they can start in a safer way from the intermediate step shown in the figure. Access Nodes providing traditional TDM subscriber concentration functions, over e.g. standard V5.2 interfaces, and with the capability to be enhanced with VoIP Gateway functionality, will prove to be an invaluable tool to facilitate the migration. In our opinion, the last equipments to be migrated to VoIP should be the ones in customer premises, due to the economic impact of their volumes. Again, an intermediate node like the Access Gateway we describe in the next section, will help in making a business sensible transition.

4 An Integrated Multiservice Access Platform with VoIP Gateway embedded

Digital Line Concentrators (DLC) have been around since long ago as a mean to extending the reach of Local Exchanges in such a way that they can cover larger areas of service, in particular, when covering disseminated population distributions. In principle, a DLC is a remote switching unit that is

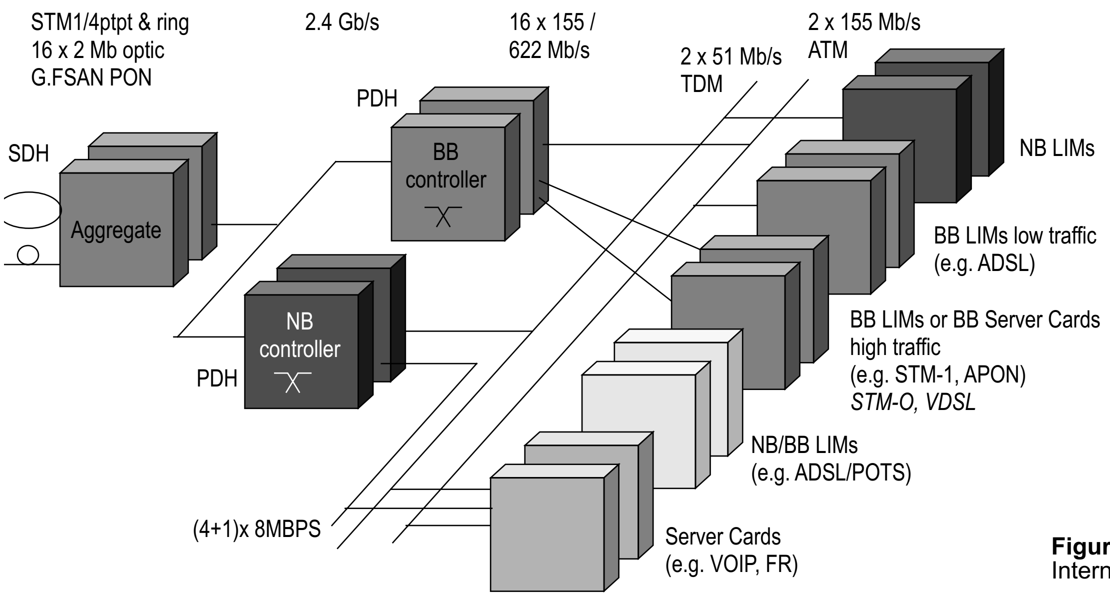


Figure 7:
Internal architecture of IMAP

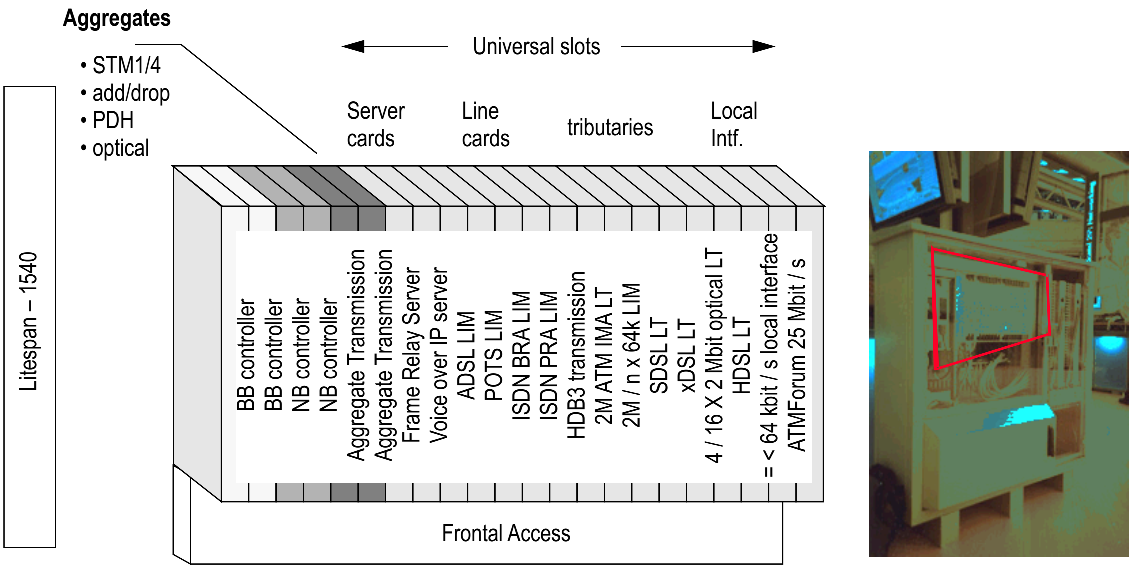


Figure 8: Physical aspect of the IMAP implementation

controlled from a Local Exchange, and that provides either multiplexing or better concentration of telephony calls for saving transmission resources between the remote location and the exchange.

DLCs have evolved in some important ways:

- Providing standard signalling interfaces to the Local Exchange (e.g. GR-303, V5.1/2), that has made independent the access nodes from the switches, and as a consequence opening the access market.
- Integrating different transmission technologies towards the exchange (copper, fibre, PDH, SDH), i.e. the same concentrating electronic node can incorporate whatever transmission technology is best suited to the deployment scenario.
- Providing a full set of user interfaces for supporting different services: from traditional telephony service (POTS), narrow band digital integrated services (ISDN), to any flavour of leased lines for narrow band data: data n x 64kbps, subrates, analogue data, etc.
- Providing the grooming of telephony and data services, i.e. the separation (network direction) of these services to the corresponding overlay network.
- Integrating Broadband Access: ADSL, HDSL, SDSL, … xDSL, which constitutes the major evolutionary step in these new generation access elements.

This conforms what is known as the New Generation Digital Loop Concentrators, or as we name it in this paper, an Integrated Multiservice Access Platform (IMAP). Figure 6 shows schematically this IMAP definition (see also reference [González Vidal 00]).

The internal architecture of an IMAP will support the following design principles, in order to provide an Operator with the necessary flexibility for its network application:

- Universal shelf and Universal slots, for subscriber and network sides: for providing the required flexibility when planning and dimensioning the access network, as the penetration of each type of service is very difficult to predict, and varies from application to application.
- Fully redundant: carrier grade, compliant with the stringent reliability requirements that public telecommunication network adheres to.
- Designed for in/outdoor configurations
- Standard compliance: in both signalling and management, to provide the basis for an open market in the access

In the platform we have developed, we have made provision for server cards, that can be located in a universal slot position. These server cards are based on digital signal processing technology and are programmed to provide VoIP, FoIP (Fax over IP), voice band data over IP. The server can also be programmed to work as a Modem Bank for providing dial-up IP service for analogue modems.

The internal architecture of the platform under discussion can be seen in Figure 7. Figure 8 shows the physical aspect of the implementation.

In our approach, TDM calls are terminated in the server board, where the bearer plane conversion takes place from 64kbps PCM channels to the voice coding scheme chosen in the IP network. Then, the voice is packetized and sent over RTP/ IP/AAL5/ATM. Packetized voice over ATM is multiplexed with data on ATM coming, for instance, from the ADSL Line Modules. ATM has been chosen as the best way to provide QoS in the access domain, and providing service multiplexing.

As for the control plane, subscriber signalling is terminated in the narrow band controlled and converted to access signalling (V5). The server card provides the interworking between this TDM signalling and the H.248 signalling towards the Call Server.

One important feature of this approach is that we can provide concurrently VoIP and traditional TDM services, by provision-

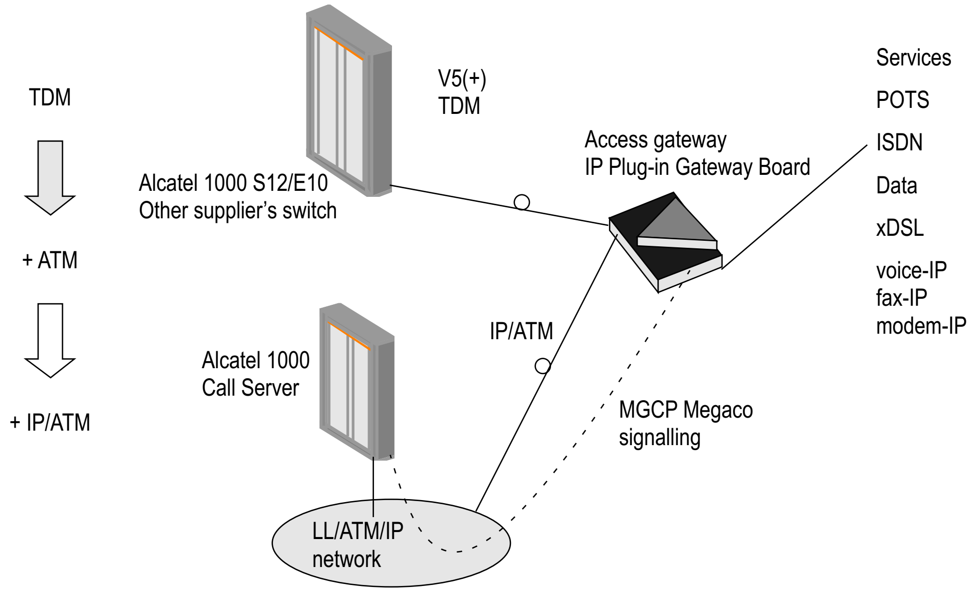


Figure 9:
IMAP provides concurrently VoIP and conventional TDM services

ing those subscribers we can serve with either technology. This feature provides a safe evolution path for operators migrating towards a voice packet network. We sketch this feature in Figure 9.

5 Conclusions

Quoting a high rank industry guru, *"The only thing we can be sure about Telecommunications future is that we cannot predict it"* we believe that there is a strong trend towards a packetized multiservice network. Most likely, this network will be based on Internet Protocol. the challenge is to replace the more-than-a-century-old telephony service.

In this paper we have briefly discussed some of the issues we are confronted to:

- what are the requirements that we have to put on IP networks for conveying voice with the adequate quality?
- what are the present models the industry is working on?
- what is the shape this Next Generation Network is going to have and what are its basic constituent elements? and
- what can be the evolution path from the current networks to the future ones?

Finally, we have described a network element that can play a major role in this transition period, that, for sure, will be long lasting, and that, in our view, it is a crucial piece in making this evolution economically sensible.

References

[ITU-T G.107]
ITU-T Recommendation G.107, The E-model, a computational model for use in transmission planning.

[De Vleeschauwer et al. 00]
D. De Vleeschauwer, J. Janssen, G.H. Petit and F. Poppe, "Quality Bounds for Packetized Voice Transport", Alcatel Telecom Review, 1st Q2000.

[De Vleeschauwer et al. 99]
D. De Vleeschauwer, J. Janssen, G.H. Petit, "Voice over IP in Access Networks", Proceedings of the Seventh IFIP Workshop on Performance Modelling and Evaluation of ATM/IP Networks, Antwerp (Belgium), 28-30 June 1999.

[González Vidal 00]
F. González Vidal, "Plataforma de Acceso Multiservicio: un seguro de la inversión frente a las incertidumbres de evolución del mercado de las Telecomunicaciones", Infrastructure & Application Summit, ITU Telecom Americas 2000.

Acknowledgements

Figures 1, 2, 3, and Table 1 are courtesy of the Alcatel Telecommunications review.

to Contents, page 1

Voice Communication over the Data Network Convergence of Services by LAN Telephony

Robert Bertels

In the past, voice and data traffic were transmitted over completely separated systems. Today, voice communication can be carried on an local data network by means of the Internet Protocol (Voice over IP or LAN Telephony). As a result, in new buildings at least, a telephone network as such is no longer necessary.

Keywords: VoIP, Internet Telephony, Convergence of Communication Services, Integrated Business Solutions, Quality of Service.

In today's office, access to the local data network is as important and normal as the telephone connection. On both networks, information is transmitted in digital form. Yet, in most cases, voice and data are carried over entirely distinct systems, a relic of the time when telephony used analogue signal transmission and computers were at an early development stage.

Voice and data on the same network

Current developments mean that the separation between voice and data transmission is no longer necessary inside a company. Local Area Network (LAN) telephony (better: Intranet telephony or VoIP) is a reality. The telephone receiver can be connected to a desktop PC and calls between offices can be made over the local data network, with no telephone line involved. Voice and data service have converged with both using the same network, opening up a basis for "Integrated Business Solutions".

Until recently, rapid progress with VoIP was predicted, along with an early end to conventional telephony. In the light of real experience, such euphoria has given way to a more realistic assessment: VoIP does not yet offer the same mix of functions and services as conventional telephony (except at very great expense) and the technical standards indispensable for a breakthrough at the global level are not yet in place. There is also a lack of know-how and experience among many vendors and potential customers. In addition, the development of VoIP is affected by the cultural chasm that exists between telephone and computer worlds as demonstrated by the fact that responsibility for voice and data transmission is divided between different departments in many companies. Despite this, the potential of VoIP should not be underestimated.

In new office buildings it can already make sense to dispense with the installation of a telephone system with exchanges and lines and use the internal data network instead for the transmission of internal telephone calls. For outside calls, a simple interface (gateway) is sufficient to provide a telephone service. At the new headquarters of *Bison Switzerland AG* (a computer company, at the time *Agro Data AG*) in Sursee (Switzerland), this new technology has been installed. Each work station has been equipped with a PC that also serves as a telephone. *Bison* adopted this revolutionary approach because of the obvious advantages of VoIP, namely:

- simple installation and administration,
- fast installation of new system features as well as easy upgrading (only software changes are required),
- open interfaces to other systems and applications (e.g. customer administration).

In existing buildings with conventional telephone systems not yet fully depreciated, total conversion is not essential and a more gradual transition to the new technology to meet current priorities is recommended. A typical example is the publisher of the *Neue Zürcher Zeitung*. The company's customer care centre was equipped with an LAN-supported communications system at the end of 1999. Calls from the subscribers are now directed into the local data network and answered from computer terminals on which the customer's data is immediately available on screen. Because it was possible to integrate the VoIP system into the existing workflow software, investment costs were relatively low. The advantages show on two levels:

- improved efficiency and quality of service, fewer customer complaints and subscription cancellations.
- from the operational point of view, the improved work environment means lower employee turnover with a consequent reduction in training costs and less need for temporary workers.

As both factors have such positive economic effects, the system will pay for itself inside a year.

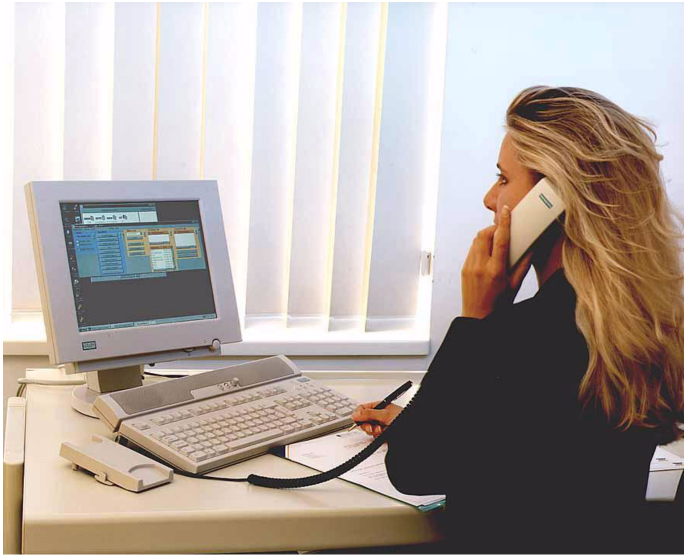
In simple terms, the model can be described as follows: the conventional local telephone system with its switchboard, tele-

Robert Bertels is head of the Venture Group for VoIP & Call Centre of Siemens Switzerland AG. He is responsible for the sales of the VoIP and Call Centre activities in the area Siemens Switzerland Enterprise networks. <robert.bertels@siemens.ch>

Siemens has been operating in Switzerland for more than hundred years. Its main area of activity is communication where its research and production function employs more than thousand people. It is thus one of the largest development centres in Switzerland. The business unit, Enterprise Networks, develops concepts and supports customers' technical installations.

to Contents, page 1

phone lines and the usual multifunctional telephone hand sets remains in operation throughout the office, except for the customer care centre. The local data network between individual workstations is also in place. In the customer care centre (and only there), the conventional telephone infrastructure is missing. PC workstations are equipped with a telephone receiver and the corresponding software to support the voice transmission. The interface between the two networks is a number of special gateways. The end result is that, after the change, there appears to be no difference in "telephone" usage and quality compared with the previous situation.



Strategic project

Companies considering the possibility of a gradual change to VoIP would be well advised to treat such a project as a high priority strategic one. Migration paths and speed must be optimised to ensure that the company's business processes are fully supported in each phase of the change if a positive added value is to be achieved. The choice of the system vendor is of particular importance. The vendor must be competent in both data transmission and conventional telephony to guarantee that the existing telephone system remains an effective element of the overall system during a possibly long transition period. All this involves concern for operational matters and the need to protect existing investment. In Switzerland, many companies have very modern and efficient telephone systems which, with the necessary expertise, can easily remain in operation until the initial investment has been paid back.

The problem

Conventional telephony is based upon the transmission of voice data over a dedicated line. In contrast, data transmission over the LAN is more flexible but also more difficult to control.

On the LAN, information is chopped in small packets which may reach their destination via different paths. If the -network becomes overloaded, packets might be delayed and arrive out of sequence. Packets can also get lost or destroyed and must be resent until they have all reached their destination. Once there, they are then recombined in the correct sequence. This process naturally takes time.

In the case of pure data this is not a problem. Voice data, however, needs to be transmitted instantaneously. There should be no perceptible time difference between what is spoken at one end of the line and what is heard at the other. such instantaneous transmission was not part of the original LAN specification but effort is now being made to achieve such a level of service. This includes greater bandwidth and its effective management, redundancy to obtain lower failure rates, and prioritization of voice data with the aim of guaranteeing a full telephone quality of service. VoIP has already been successfully implemented in many company Intranets, and is well on the way to becoming the generally accepted solution in the Wide Area Network field too.

References

[Davidson/Peters 00]
Jonathan Davidson und James Peters. Voice Over IP. Grundlagen, Konzepte, Technologien, Anwendungen. Cisco Press, 2000 (ISBN 3827258006)

[Douskalis 99]
Bill Douskalis. IP Telephony. The Integration of Robust VoIP Services. Prentice Hall, 12/99 (ISBN 0-130141186).

[Foth 01]
Egmont Foth. Handbuch IP-Telefonie. Fossil Verlag, Köln, 2001 (ISBN 3-931959-33-3).

[Goncalves 98]
Marcus Goncalves. Voice over IP Networks. McGraw Hill, 1998 (ISBN: 0-07-913783).

[Held 98]
Gilbert Held. Voice over Data Networks. Covering IP and Frame Relay. McGraw Hill, 1998 (ISBN: 0-07-028135-1).

[Karapetkov 99]
Stefan Karapetkov. Siemens Information and Broadband Networks. November 30, 1999

[Minoli 98]
Daniel Minoli. Delivering Voice Over IP Networks. Wiley Canada, 1998 (ISBN: 0-471-25482-7).

to Contents, page 1

DAMAGED WATERMARKS DETECTION IN FREQUENCY DOMAIN AS A PRIMARY METHOD FOR VIDEO CONCEALMENT

Robert HUDEC[1], *Martin BOJMIR*[1], *Miroslav BENCO*[1]

[1]Department of Telecommunications and Multimedia, Faculty of Electrical Engineering, University of Zilina, Univerzitna 8215/1, 010 26 Zilina, Slovakia

robert.hudec@fel.uniza.sk, martin.bojmir@fel.uniza.sk, miroslav.benco@fel.uniza.sk

Abstract. *This paper deals with video transmission over lossy communication networks. The main idea is to develop video concealment method for information losses and errors correction. At the beginning, three main groups of video concealment methods, divided by encoder/decoder collaboration, are briefly described. The modified algorithm based on the detection and filtration of damaged watermark blocks encapsulated to the transmitted video was developed. Finally, the efficiency of developed algorithm is presented in experimental part of this paper.*

Keywords

Video concealment, watermark, lossy channel, video quality, temporal filtration, SSIM.

1. Introduction

The main problems arising by the video transmission over IP communication networks are caused by network congestion. With growing of the multimedia services popularity, video concealment techniques become more important. A lot of video concealment methods with different efficiency have been proposed. They can be divided into three primary groups, namely, forward error concealment, interactive video concealment and post-processing video concealment. These methods were developed to work separately. Moreover, in this paper, the new combined concealment algorithm for increasing the efficiency of error cancellation was developed [1].

The outline of the paper is as follows. In the next two sections, an overview of basic principles of video concealment techniques, especially the post/processing methods are introduced. Based on state of art in this area, the new method utilizes the damaged watermark detection algorithm is presented in the section 4. Finally, the realized experiments, brief summary and future tasks are discussed in sections 5 and 6.

2. Methods of Video Concealment

In recent years, three basic types of video concealment have been developed. These approaches include forward, interactive and post-processing video concealment. In first type of video concealment, the encoder plays leading role. If transmission channel is not errorless, the two kinds of distortions in decoder can be observed. First one, quantization noise caused by waveform coder and second one, caused by transmission error. Optimal pair of source and channel encoder should minimize both types of distortion. Thus, it can achieve better performance of video transmission. Forward video concealment includes FEC (Forward Error Correction), common application of source and channel coder, as well as layered coding transmitted by channels with different priority. Some of these methods require the cooperation of network and coders in order to achieve different level of quality of service for diverse parts of video stream. However, the base layer contains principal video information, it is transferred with higher reliability. The higher video quality can be achieved by additional video information layer enhancement. Design of layered coding can be solved by several approaches, for example in temporal domain with different frame rates or in spatial domain with different resolutions. Most important task is partition of video data to given layers. The transmission error in enhancement layer could decrease of video quality. This method of video concealment adds redundancy into source or channel coder and also utilizes transport prioritization. The main advantage of this method is good error resistance due to increase of the overhead.

Forward video concealment and post-processing video concealment utilize a very small interaction

to Contents, page 1

between encoder and decoder. If a backward channel from decoder to encoder is available, these devices can optimally cooperate and achieve better results by the damaged video processing. This cooperation can be realized by source or channel coder. Source coder can adapt coding parameters based on the backward channel information from decoder. Also the information from backward channel for the bandwidth reservation for FEC or repeated transmission can be used. There are also other techniques of the interactive video concealment, for example; selective coding in source encoder, adaptive coding on channel level, resending lost data or damaged parts without waiting, as well as prioritization and multiple copies of resending data. The decoder sends information about correctly received data. Encoder from this information can determine damaged parts. Subsequently, these parts are replaced with correct parts from encoder buffer and used for prediction. This method helps to reach higher error resilience. Automatic repeat request can be used together with conventional decoder. In general, one retransmission creates delay about 70 ms. This value is appropriate for most of non-real time applications. On the other hand, it causes delay in decoding that may be unacceptable for real-time applications. On suppression problem of delay in transmission, there has been proposed a method that averts the delay in decoding by remembering the path of damaged data at decoder side. Other method decreases delay in decoding by sending multiple copies of lost data in every repeated transmission due to reduction of the number of requested repeated transmissions.

The last one video concealment method uses spatial and temporal redundancy presented in video signal. The filter dimension used for the interpolation the lossy blocks strictly depends on the on the amount of motion in the concealed area of processed frame. There are some combinations of temporal and spatial causal interpolation masks. In the case of using the frame memory, the interpolation masks can have a non-causal form. Thus, the temporal causal concealment techniques repairs lost blocks with corresponding blocks from previous frames, spatial concealment techniques calculate lost pixels from neighborhood pixels just the same frame and finally, the spatio-temporal concealment techniques merge previous mentioned principles. Moreover, the spatial interpolation techniques are preferred to use for video with low amount of motion. On the other hand, the temporal interpolation techniques are preferred for video with high amount of motion [1], [2].

3. Post-Processing Video Concealment

In this section, some of post-processing video concealment methods will be introduced. They are characterized by minimal interaction between encoder and decoder and use correlation between temporal and spatial pixels/blocks. Other ones use estimation of optical flow at pixel or block level and extrapolation of motion vectors. Following methods repair whole damage frames by multiframe video concealment in order to make effective concealment of errors generated during the transmission over network.

3.1. Hybrid Method of Video Concealment

The hybrid method utilizes spatial and temporal redundancy in video signal. This method is trying to mask lost blocks by using the information from two previous frames and spatial adjacent blocks of the actual frame. Moreover, spatial concealment uses directly adjacent vectors. These vectors are also use for the smooth estimation of lost blocks. Temporal error concealment conceals lost block with block from previous frame. Further, in this method the PDC (Pixel Difference Classification) function is used. PDC function compares pixels each other in the processed frame and pixels at the same position in previous frame. PDC function detects motion in video. When the value of PDC function is higher than threshold, the algorithm uses spatial concealment only. If detected motion in video is too much for temporal concealment, so it might make inaccurate concealment. This algorithm can be used for application that does not request very high video quality such as video conference or broadcasting for mobile phones [2], [3].

3.2. Video Concealment Based on Boundary Distortion Estimation

This hybrid video concealment algorithm utilizes not only different types of concealment techniques but combine whole schemes. Temporal video concealment provides better performance than spatial concealment on general conditions, but spatial concealment is more convenient for video with dynamic scene. Application of these techniques depends on various parameters such as temporal and spatial activity. Furthermore, concealment algorithm uses enhanced techniques in order to achieve more accuracy recovery of lost parts of video frames. Temporal concealment contains two kinds of recovery of lost video data. Convention temporal method looks for the most similar MBs. Second method is performed by addition of enhanced residuals to recovery MBs based on calculate of boundary residuals. Then TAWS (Temporal Adaptive Weight-based Switching) algorithm that processes data from both types of concealment to finally form of MBs is accomplished. Spatial concealment is based on SAWS (Spatial Adaptive Weight-based Switching) algorithm that determines suitable type of spatial concealment. A few techniques such as search of dominant edge points, integration of interpolation modes

to Contents, page 1

in order to obtain more accuracy performance are utilized. Usage of convention techniques of concealment and enhanced algorithms such as TAWS, SAWS and boundary distortion estimation leads to very good performance [4].

3.3. Video Concealment Algorithm on Pixel-C_{Ap}

Video concealment usually uses the information from adjacent blocks near lost block. In case that whole frame is lost, conventional concealment methods fail because the information from adjacent blocks is unavailable. Thus, in these cases other methods for the motion estimation from previous decoded frames in order to conceal the lost frame have to be applied. The method based on pixels has been proposed in situation, when whole frame is lost, so this method is based on the optical flow estimation. Optical flow estimation is based on the hypothesis, that motion between two following frames is not very different. This technique of concealment of lost frame is based on projecting adjacent correctly decoded frame on lost frame in pixel domain. Motion in video can be assessed by estimation of motion vector field. For properly and effective concealment of lost frame is necessary perform several steps: estimation of motion vector field, their spatial adjustment, projecting on lost frame and finally, interpolation and filtration [5].

3.4. Video Concealment Algorithm on Block-C_{Ab}

Video concealment method based on block level tries to provide a concealment that will be possible to implement in real-time applications. CAb method utilizes the concept of optical flow to reconstruct field of motion vectors for the lost frame [5]. Reconstruction on block level offers two main advantages:

- algorithm is used in case, when decoder detects random lost in picture. Solution is assumed by single filling of the data structure used with normal macroblock decoding and data filtering. Therefore this form can be used in parallel architecture, where filtration and interpolation are processed together,
- all filtration operations are needed for interpolation of lost information, for example motion vectors from intra-frame coding areas, based on block level with resolution 4x4 pixels. This resolution enables more accurate processing than C_{Ap} method.

This concealment algorithm for loss frame on block domain is performed gradually in several steps. At first, in order to increase coding effectiveness the convenient reference frame is searched. It means that algorithm searches the closest frame in time in reference buffer. Projection on motion vectors of all pixels from reference frame is exploited and applied on lost frame, subsequently. In next step, statistical value for moving vectors in macroblocks and blocks with different size is computed. In order to achieve more precision assess of motion, estimation for macroblock and block levels is performed. Algorithm tries to determine more uniform motion in image on macroblock level such as background and for remaining part of image is used estimation on block level.

3.5. Video Concealment with HMVE

Motion vectors of lost pixels that are extrapolated from previous frame motion vectors may be not accurate. Some of them could by badly extrapolated, especially in video with very high amount of motion. These limitations decrease the accuracy of motion vector for a pixel as well as total performance. Owing to this fact some techniques have been developed. These techniques utilize extended extrapolation of motion vectors based on pixel level. In order to remove these limitations there has been proposed a hybrid motion vector extrapolation method based on PMVE (Pixel Motion Vector Extrapolation), that uses not only extrapolation motion vectors of pixels, but it also extrapolates motion vectors of blocks too. This method is capable to eliminate badly extrapolated motion vectors in order to achieve accurate motion vectors. Moreover, HMVE (Hybrid Motion Vector Extrapolation) video concealment method works with block resolution 4x4 pixels and video standard H.264/AVC that uses compensation blocks with resolution 4x4 pixels too [6], [7].

3.6. Multiframe Concealment Method

This method uses not only previous decoded frames, but also information from partly decoded following frames in order to increase quality of the lost frame and also quality of following frames. Macroblocks of the lost frame are concealed in convenient order that is specified by using the information from partly decoded following frames. This algorithm can be adjusted for different number of reference frames that are used for successful concealment of damage parts of frames. High number of used frame can lead to more accurate concealment, but on the other hand high number of frames may not be suitable for real-time applications. Operations of concealment are performed on two levels - macroblock domain and also on blocks with resolution 4x4 pixels. In first step, algorithm decodes following frames from damaged frame, consequently determines priority value of decoded macroblocks in damaged frame and adjusts its appropriate order. Macroblocks in damage frame are concealed by suitable motion vectors. After concealment, residual parts of following frame can be correctly

to Contents, page 1

decoded.

Errors in received frame not only decrease the quality of video sequence, but also cause error propagation in following frames. Video concealment may minimize the immediate impact of lost packet on actual frame and may minimize error propagation. Multiframe concealment based methods provide better error concealment than previous methods due to objective and subjective quality of lost frames concealment and minimize the error propagation [8]. Special kind of multiframe concealment was applied on multi-view 3D video sequences where the multiple frames captured by cameras contain the same scene from different viewpoints [13].

4. Watermark Based Concealment Method in Frequency Domain

In this section, the new algorithm for the detecting the lost video blocks inside the video transmission is developed. It is based on the well known method for measuring the video quality, namely, the watermark encapsulation method. The entire algorithm is divided to the 3 main parts, namely, watermark embedding process at the transmitter's side, lost block detector and lost block of pixels interpolation at the receiver's side as shown in Fig. 1. Moreover, the receiver has to know the watermark embedded to transmitted video [12].

Fig. 1: Principal block diagram of developed method.

4.1. Watermark Embedding

At the beginning, the choice of watermark embedding technique to video with minimum affect to video quality has to be chosen. There are several options how the watermark can be inserted into video, namely, in spatial or frequency domain and likewise, into luminance or chrominance components. Based on previous research, the embedding in frequency domain into luminance component was selected. The watermarks serve as some kind of feature to measuring the error transmissions over the network.

In first order, to the each frame of transmitted video the binary watermark in 8x8 size dimensions to each DCT block of luminance part of YCbCr color space video in frequency domain have been embedded. Higher human eye sensitivity on luminance and the simplicity of whole algorithm are the main reasons (Fig. 2).

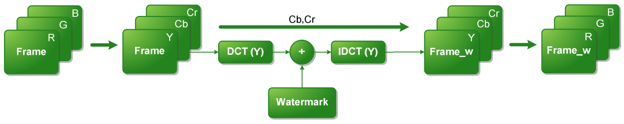


Fig. 2: Principal block diagram of developed method.

4.2. Detector of Damaged Watermark in DCT Block

Let the receiver transforms the received frame of video from RGB or YUV to YC_bC_r or receives directly Y component of YC_bC_r color space. Thus, the main role of error detector is correctly detects of damaged DCT blocks transmitted via lossy channel by known watermark.

Let the inserted watermark w has the same dimension as DCT block and is composed by bipolar values belong to the high (HF), middle (MF) or lower (LF) image frequencies (Fig. 3a). Best choice for watermark embedding to DCT blocks is using the MF area (Fig. 3b). Moreover, the high values of LF coefficients can cause the periodic appear of white pixel in left-upper corner in time domain and high values of HF coefficients will cause visible level of Gaussian noise.

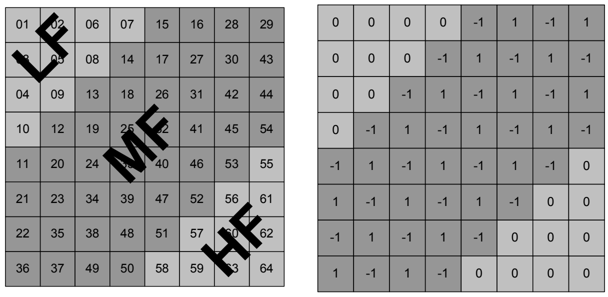


Fig. 3: DCT block, a) dividing to image frequencies, b) an example of bipolar watermark "Chess" concentrated to MF area.

In first step, the masks of positive and negative values of the watermark's DCT coefficients are computed as follows:

$$mask^{+}(i,j) = LOG1\{DCT(i,j) = MAX\{DCT(i,j)\}\}, \quad mask^{-}(i,j) = LOG1\{DCT(i,j) = MIN\{DCT(i,j)\}\} \tag{1}$$

where the LOG1{.} express the logical function for true value (logical 1) of expression located in the brackets and i, j are the coordinates of DCT coefficient. The Likewise, the overall numbers of positive and negative values can be computed too as follows

$$w^{+} = \sum_{i=1}^{8}\sum_{j=1}^{8} mask^{+}(i,j), \quad w^{-} = \sum_{i=1}^{8}\sum_{j=1}^{8} mask^{-}(i,j). \tag{2}$$

Let the $Threshold_{UP}$ and $Threshold_{DOWN}$ set up the acceptance level of undamaged value of watermarked image and each element of detected results can be

to Contents, page 1

computed as follows:

$$d^+(i,j) = mask^+(i,j) * \{DCT(i,j) > Threslold_{UP}\}$$
$$d^-(i,j) = mask^-(i,j) * \{DCT(i,j) < Threshold_{DOWN}\} \quad .(3)$$

The overall decision if DCT block was corrupted or not is classified by next formulas:

$$decision = \frac{\sum_{i=1}^{8}\sum_{j=1}^{8} d^+(i,j) + d^-(i,j)}{w^+ + w^-}, \quad (4)$$

$$Lost_{DCT}(Y) = LOG1\{decision > Threshold_{SCORE}\}, \quad (5)$$

where Y is luminance channel of watermarked image received by receiver. In case, where the algorithm evaluates some block as damaged, block Y and appropriate C_b and C_r color blocks are marked in mask of lost DCT blocks $Lost_{DCT}$ (Y) as value 1 (True). Creation of lost DCT blocks mask inside one video frame mitigates realization of error concealment mechanisms on such damaged positions [9].

4.3. Lost DCT Block Interpolation

In the case of correct identification of lost DCT block using detection of damaged watermark in luminance channel, a lot of interpolation techniques can be used. In this paper, the simple time VMF (Vector Median Filter) belongs to the order-statistics filter family is used [11].

5. Experimental Results

In the experimental part, the testing video sequences to indentify the algorithm performance were used.

5.1. Test Conditions

At first, the two static color images called Lena (representative of simple image) and Mandrill (representative of detailed image) as test image data with resolution 256x256 pixels and contain 1024 DCT blocks by 8x8 coefficients were used. Moreover, they are shown in Fig. 4a. These reference images were tested for binary watermark pattern in order to apply on video sequence. Thus in Fig. 4c, the test images with embedded watermark pattern chess <1;-1> are shown. Moreover, watermarked image Lena achieved SSIM parameter equal to 0,9876 and Mandrill 0,9957.

A lot of types of watermarks were tested. At first, watermarks have been inserted into all DCT coefficients. In these experiments, bipolar binary watermarks called chessboard, horizontal lines, vertical lines, text with combination of values: <0;1>, <1;-1> and their multiplications were tested. They are shown in Fig. 4b.

In the Tab. 1, the impact of used watermark pattern on number of false detected block for $Threshold_{SCORE}$ = 0,4 is presented. Experiment shows that pattern "Chess" with binary range at <1; -1> achieved 10 false detected DCT block only. More positive is fact that better results are obtained for detailed image Mandrill.

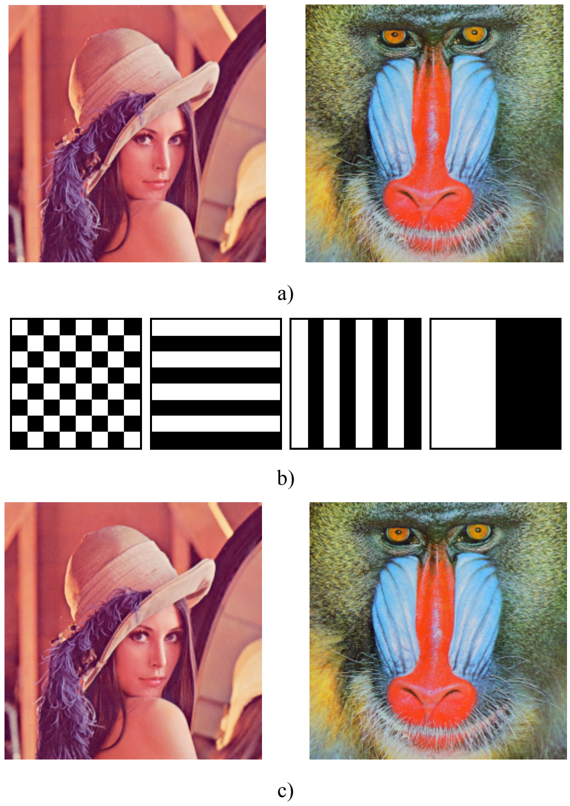


Fig. 4: a) tested watermarks with different B&W patterns, b) tested images for embedding watermarks "Lena and Mandrill", c) test images with embedded watermark "Chess, <1, -1>".

Embedded watermark for all coefficients produced image artifacts mentioned in previous section. The choice of watermark strongly depends on two opposite scopes, namely, the video quality maximization and ability to detect the damages in video. This compromise leads to using the chessboard watermark in MF DCT coefficients with <1,-1> values. Based on adjusted appropriate threshold score achieved on the test color images, the most suitable watermark were encapsulated into video sequences.

Tab.1: Impact of watermark pattern on number of false detected DCT blocks, $THRESHOLD_{SCORE}$=0,4.

Watermark pattern	**False detected blocks**		**Σ**	**Rank**
	Lena	Mandrill		
Chess, <1,-1>	7	3	10	1.
Vertical lines, <1,-1>	7	12	19	4.
Horizontal lines, <1,-1>	7	6	13	2.

to Contents, page 1

Half by half, <1,-1>	6	17	23	5.
Chess, <0,1>	16	9	25	6.
Vertical lines, <0,1>	7	20	27	7.
Horizontal lines, <0,1>	5	13	18	3.
Half by half, <0,1>	13	27	40	8.

Following experiments with static color images and two test video sequences, namely „Taxi“ as delegate of real video (classic movie) and „IceAge“ as delegate of artificial video (animated cartoon) were used too. The Taxi video is about the discussion of two people at a car and IceAge video is about the discussion of animated animals. Both videos had same dimension 320x120 pxl with 25 fps and 250 frames in total duration 10 seconds coded in Indeo5 codec (IV50) supported by MATLAB.

In the Fig. 5, the false detection of DCT blocks (8x8 pixels) with embedded watermarks vs. threshold score are shown. Moreover, some chessboard watermarks with various parameters were tested too. The experiment shows, that more suitable value of threshold laid under threshold 0,4.

5.2. Lossy Channel Modelling

The algorithm was tested as offline application, where the errors caused by video transmission via lossy communication channel were modelled. There were various types of errors. Moreover, approach of the errors modeling by uniform and Gaussian PDF functions were applied. The achieved results were measured by well known parameter SSIM (Structural SIMilarity). The test videos at size 320x120 pxl had overall number of 8x8 DCT blocks equal to 600 in one frame (40 in row by 15 in column).

Tab.2: Percentage expression of lost DCT blocks in one video frame.

PDF		uniform	normal
0,01	No.	3	0
	%	0,5	0
0,05	No.	17	13
	%	2,8	2,52
0,1	No.	30	47
	%	5	7,8
0,15	No.	48	58
	%	8	9,7
0,2	No.	55	83
	%	9,2	13,8

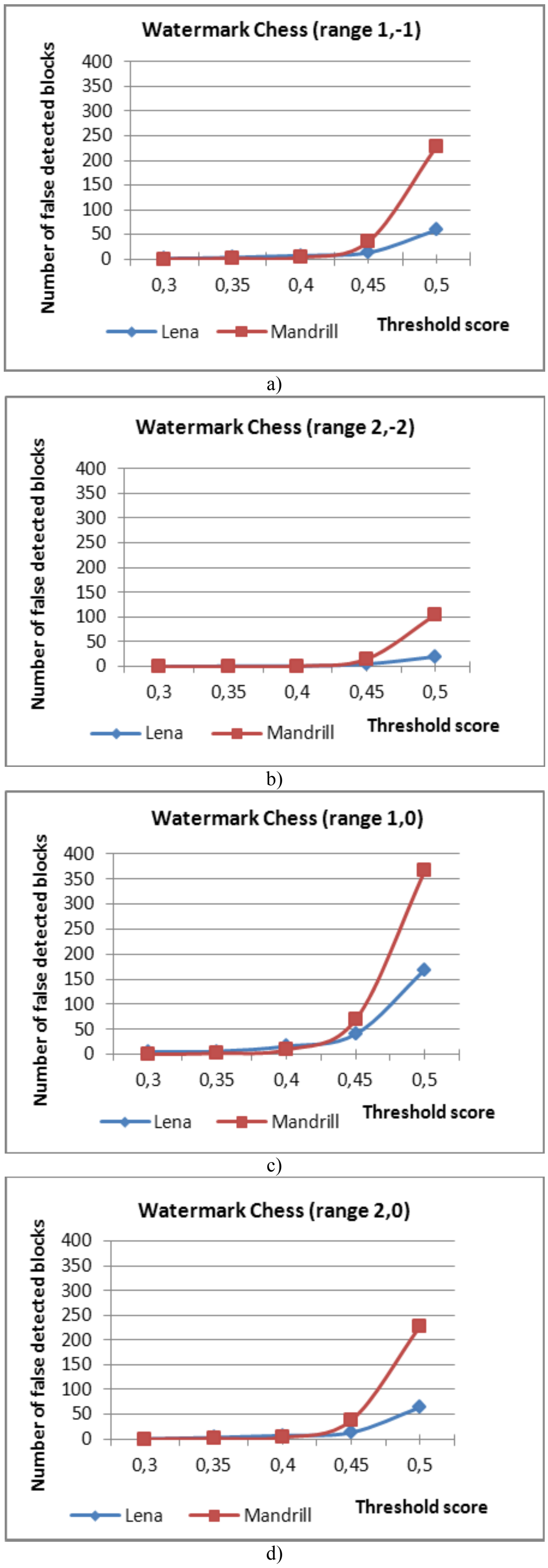
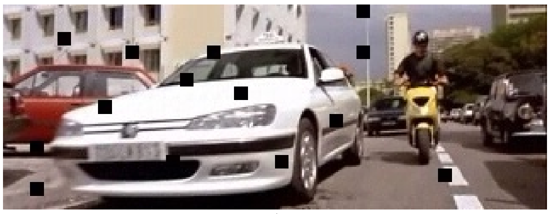


Fig. 5: False detected blocks dependency on $\text{Threshold}_{\text{SCORE}}$, a) watermark chess <-1,1>, b) watermark chess <-2,2>, c) watermark chess <1,0>, d) watermark chess <2,0>.

to Contents, page 1

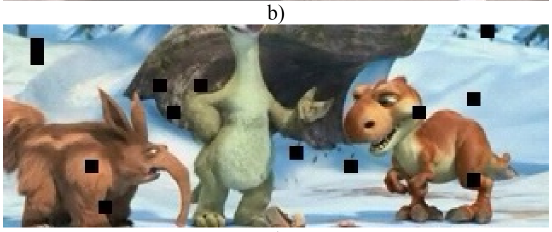
5.3. Developed Algorithm Behaviour and Results

It is very important, that after process of inserting watermark into video, the average video quality according to objective SSIM value may not decrease under 0,98. Based on this condition, the average value of SSIM parameter for embedded watermark was not lower than 0,99 [9-10]. After embedding the watermarks into test videos, achieved SSIM values were at levels equal to 0,9905 for Taxi video sequence and 0,9910 for IceAge video sequence [12].

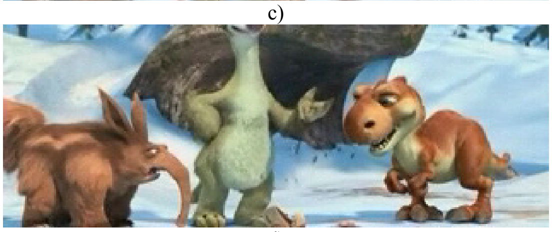


a)

b)



c)

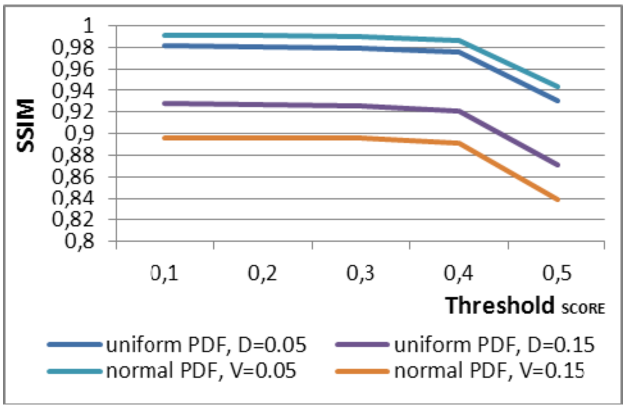


d)

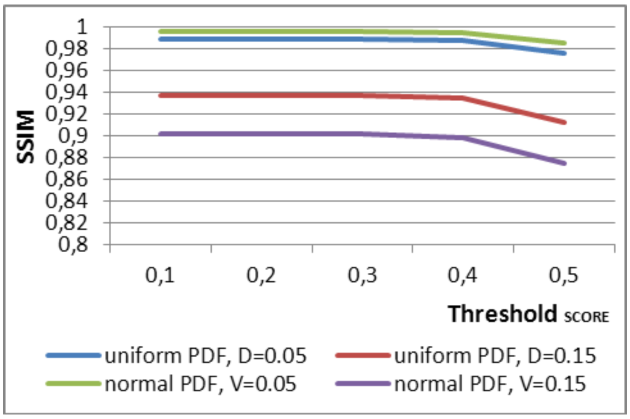
Fig. 6: Frame of test video for normal PDF and *V=0,05*, a) damaged video “Taxi”, b) concealed video “Taxi”, c) damaged video “IceAge”, d) concealed video “IceAge”.

For correct detection of corrupted watermarks, the thresholds have to be adjusted. Thus, these thresholds determine the detection correctness. If the calculated value is lower than threshold, then algorithm evaluates processed DCT block as damaged and initializes the interpolation process.

As the interpolation method, the VMF (Vector Median Filter) with time filtration mask was used. It is very powerful method in the cases, when the restorations of fast changed video data are required. Moreover, its vector approach affects interpolation with low level of CD (Color Difference). In the terms of DCT block loosing, the usage of time filtering over the spatial one is preferred. The main reason is low interpolation error in centre of DCT blocks. Because the probability of damaged DCT blocks at the same position for some following video frames is low, the VMF may not to use long filter window. In experiments, the filter window with length equal to 5 was used. Thus, the short filter window was produce fast interpolation. Moreover, in experiments, the non-causal VMF, namely, two previous and two following frames according to present frame based on the measuring the Euclidean distance (L2) were used. The filtration of DCT blocks marked as lost was realized after conversion of pixels of non-corrupted DCT blocks from YC_BC_R to RGB color space. After this, the filtration on all 64 pixels of lost DCT block was realized. The interpolated frames by time VMF are presented in the Fig. 4b and 4d (normal PDF, *V=0,03*).



a)



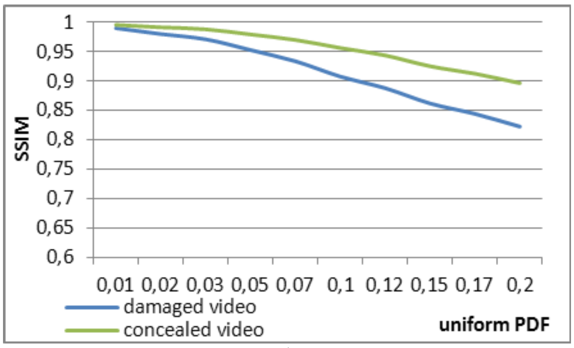
b)

Fig. 7: SSIM dependency on $\text{Threshold}_{\text{SCORE}}$ for a) real test video “Taxi”, b) real test video “IceAge”.

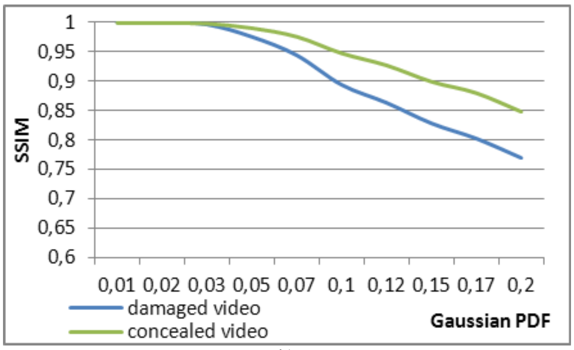
The SSIM parameter dependencies on ThresholdSCORE for test video Taxi and for two PDF parameters (Gaussian PDF) were tested. The optimal value was detected experimentally up to 0,4 for both *V=0,05* and *V=0,15* and for following experiments was used equal to 0,2 (shown in Fig. 7). The values of $\text{Threshold}_{\text{UP}}$ and $\text{Threshold}_{\text{DOWN}}$ for adjusting the acceptance level of undamaged value of watermarked

to Contents, page 1

image were in all experiments identical and equal to 0. Because the embedded watermark is bipolar <1,-1>, the equal value at zero level is statistically appropriate. Achieved curves are presented in the Fig. 7.



a)

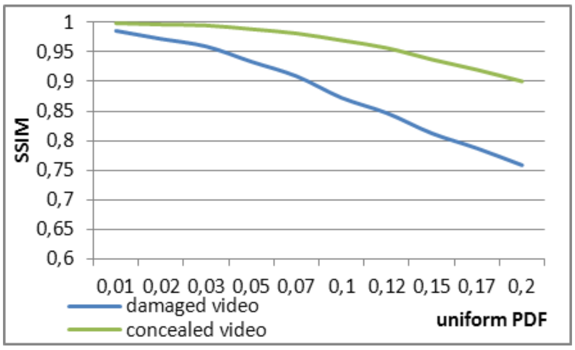


b)

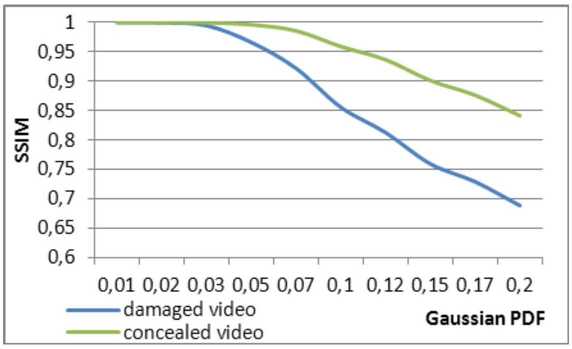
Fig. 8: SSIM dependency on PDF parameter for test video "Taxi", a) uniform PDF, b) normal PDF.

The next experiments were focused to find out the algorithm behaviors in terms of real and cartoon videos. Moreover, the dependency quality parameter SSIM on the PDF parameters was investigated. The PDF parameter for both distributions, namely, normal and uniform in range from 0,01 to 0,2 as shown in Tab. 2. was adjusted. In Fig. 7 and 8, the achieved curves for damaged and concealed videos are presented. From the experiments is evident that average values of video quality parameter SSIM is increasing. Thus it confirms that video quality after application of developed concealment algorithm was improved.

There is not absolute concealment of all damaged blocks, because the detector is not able to detect all damaged watermarks. Anyway, the developed algorithm achieves similar results for classic movie as well as for animated cartoon. Moreover, the level of damage for the test video IceAge is higher than Taxi for both PDFs. After application of the error concealment technique, both test videos at the same PDF parameter had some joint level of SSIM. Thus, for uniform PDF, it was around 0,9 and for normal PDF around 0,85. These parameters were measured at the PDF parameter D or V equal to 0,2.



a)



b)

Fig. 9: SSIM dependency on PDF parameter for test video "IceAge", a) uniform PDF, b) normal PDF.

6. Conclusion

In this paper, the novel method for video concealment have been proposed and tested. Developed algorithm is based on the detection the damaged watermarks encapsulated to the transmitted video. Thus the watermarks are encapsulated to video frames in the frequency domain. For the simulation requirements, the transmission errors via communications networks by normal and uniform distribution were used. Developed algorithm uses the detector of damaged block by identifying the lost watermark in DCT blocks. Finally, the time vector median filter for interpolation of lost blocks was used. The realized experiments approve the efficiency of developed algorithm mentioned in experimental part of this paper. In the future algorithm improvements, the temporal-spatial concealment filter masks, enhancement filters, adaptive detector of lost watermarks and shot boundary detector could be adopted. Likewise, the enhancement method for encapsulation and detection of watermarks into video frames could improve the results too. In order to testing the developed algorithm in real operation, the test video data transmission via real channel has to be used.

to Contents, page 1

Acknowledgements

This paper has been supported by the Slovak Scientist project VEGA grant agency, project no. 1/0570/10, "Algorithms research for automatic analysis of multimedia data".

References

[1] WANG, Y.; ZHU, Q. Error Control and Concealment for Video Communication: A Review. *Proceedings of the IEEE*, 1998. Vol. 86, No. 5, pp. 974 – 997. ISSN 0018-9219.

[2] MARCHEVSKY, S.; MOCHNAC, J. Hybrid Concealment Mechanism, *Acta Electrotechnica et Informatica*, 2008. Vol. 8, No. 1, pp. 11-15. ISSN 1335-8243.

[3] HRUSOVSKY, B.; MOCHNAC, J. Extended error concealment algorithm for intra-frames in H.264/AVC. *Acta Electrotechnica et Informatica*, 2010. Vol. 10, No. 4, Available at WWW: <http://www.aei.tuke.sk/pdf/2010-04/12_Hrusovsky.pdf>. pp. 59-63. ISSN 1335-8243.

[4] SHINFENG, D., L.; WANG, Ch., Ch.; CHUANG, Ch., Y.; FU, K., R. *A Hybrid Error Concealment Technique for H.264/AVC Based on Boundary Distortion Estimation*, InTech, 2011. Available at WWW: <http://www.intechopen.com/download/pdf/pdfs_id/16274>. pp. 243-258 .ISBN 978-953-307-181-7.

[5] BACCICHET, P.; BAGNI, D.; CHIMIENTI, A.; PEZZONI, L.; ROVATI, F., S. Frame Concealment for H.264/AVC Decoders. *IEEE Transaction on Consumer Electronics*. 2005, Vol. 51, Issue 1. pp. 227 – 233. ISSN 0098-3063.

[6] YAN, B.; GHARAVI, H. Efficient Error Concealment for Whole-Frame Loss Based on H.264/AVC. In *15th IEEE International Conference on Image Processing, ICIP 2008* [online]. 2008. [cit. 2011-10-17]. Available at WWW: <http://www.antd.nist.gov/pubs/ICIP_published.pdf>. pp. 3064 – 3067. ISBN 978-1-4244-1764-3.

[7] YAN, B.; GHARAVI, H. A Hybrid Frame Concealment Algorithm for H.264/AVC. In *IEEE Transaction on Image Processing*. 2010. Vol. 19, Issue 1, [cit. 2011-10-17]. pp. 98 – 107. ISSN 1057-7149.

[8] LIU, Y.; JIAJUN, B.; CHUN, CH.; LINJIAN, M.; KEWEI H. Multiframe Error Concealment for Whole-Frame Loss in H.264/AVC. In *IEEE International Conference on Image Processing, 2007. ICIP 2007.* 2007. Vol. 4, [cit. 2011-10-17]. pp. IV - 281 - IV – 284. ISBN 978-1-4244-1437-6.

[9] KEITH, J. *Video demystified: A Handbook for the digital Engineer*, Massachusetts, USA, Elsevier Inc. 2005, pp. 927. ISBN 0-7506-7822-4

[10] VACULA, M. *Digital watermarks in color images*. Kosice, 2003. Diploma thesis. Technical University of Kosice.

[11] ASTOLA, J.; HAAVISTO, P.; NEUVO, Y. Vector Median Filters, *Proceedings of the IEEE*, Vol. 78, No. 4, April 1990, pp. 678-689. ISSN 0018-9219.

[12] BOJMIR, M. *Videoconcealment: Optimization and evaluation in telecommunications networks*, Diploma thesis, Department of Telecommunications and Multimedia, FEE University of Zilina, pp. 90, 2010.

[13] HRUSOVSKY, B.; MARCHEVSKY, S.; MACEKOVA, L. Error Concealment Algorithms Applied on Multi-View Video Sequences. *Acta Electrotechnica et Informatica*, Versita, Warsaw, 2011. Vol. 11, No. 2, pp. 11-16. ISSN1335-8243.

[14] CHUNG, T.; SONG, K.; KIM, Ch.. *Error Concealment Techniques for Multi-view Video Sequences,* Lecture Notes in Computer Science, 2007, Vol. 4810/2007, pp. 619-627. ISSN 0302-9743.

About Authors

Robert HUDEC was born in Revuca in 1974, Slovakia. He received his M.Sc. and Ph.D. degrees from the Technical University in Kosice, Faculty of Electrical Engineering and Informatics, Slovakia, in 1998 and 2003, respectively. Nowadays, he is an Associated Professor at the Department of Telecommunications and Multimedia, University of Zilina. Areas of his research includes digital image processing with special focus laid on filtration of mixed noise by using adaptive order-statistics filters and combined low-level image description for MPEG-7 standard. From 2005 his research interests covers also systems, services and terminals of e-/m-health applications. He received Werner von Siemens Excellent Award and Jozef Murgas Award for his research on image filtration in 2003 and 2007, respectively, and Award of „Vice premier and minister of school for science and research“ in the category: celebrity of research and development till 35 years old for his research on e/m-heath systems and terminals in 2006.

Martin BOJMIR was born in Zilina in 1985, Slovakia. He received his M.Sc. degree in Telecommunications from University of Zilina, Faculty of Electrical Engineering, Slovakia, in 2009. Currently, he is a Ph.D. student at the same faculty. His areas of interest include video quality assessment, networking and digital image processing applied in Intelligent Transportation Systems.

Miroslav BENCO was born in 1981 in Vranov nad Toplou, Slovakia. He received his M.Sc. degree in 2005 at the Department of Control and Information Systems and Ph.D. degree in 2009 at department of Telecommunications and Multimedia, University of Zilina. Since January 2009 he is researcher at the Department of Telecommunications, University of Zilina. His research interest includes digital image processing, semantic analysis of multimedia content, image and video classification and retrieval for MPEG-7 standard.

to Contents, page 1

THE EFFICIENCY OF CONSTRAINT BASED ROUTING IN MPLS NETWORKS

Martin MEDVECKY[1]

[1]Department of Telecommunications, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, Ilkovicova 3, 812 19 Bratislava, Slovakia

martin.medvecky@stuba.sk

Abstract. *The paper presents the simulation results and evaluates the efficiency of constraint base routing algorithms used in MPLS network from the point of their usability in Next Generation Networks. The efficiency of constraint based routing is evaluated according following criteria: optimal path selection, routing priority of traffic flows selected for constraint routing and bandwidth allocation by MAM or RDM bandwidth constraints models.*

Keywords

Constraint routing, maximum allocation model, Russian Doll Model, MPLS, QoS.

1. Introduction

The MPLS (Multiprotocol Label Switching) [1] is a connection oriented packet network technology originally proposed to simplify and speed up a packet processing in broadband network nodes. Today the MPLS is considered as a core transport technology for NGN (Next Generation Network), therefore it should support real-time QoS sensitive services like voice, video, etc. Real time services have a strict QoS requirement on delay, jitter, packet loss etc [2]. There were no implicit QoS mechanisms proposed in MPLS, the quality of service can be achieved by a combination of several individual mechanisms. The proposed paper investigates the traffic routing methods as a one of the Quality of Service supporting mechanism.

There are three types of traffic routing used in MPLS:

- Hop-by-hop routing,
- Explicit routing,
- Constraint routing.

The hop-by-hop routing use a plain IP routing algorithms and does not provide adequate QoS quality due to impossibility to distinguish routing of packets through the network according to QoS constraints. The Label Switched Path (LSP) is established according the route selected by routing algorithms, usually a shortest path is selected, and in the fault state (congestion, connection breakdown, etc.) the guaranteed bandwidth is not available even for packets with highest priority.

Explicit routing uses LSPs explicitly selected by a network administrator. The routing can be considering QoS parameters and a real state of the network. The traffic can be classified into different QoS classis and for each Forwarding Equivalency Class (FEC) a different route can be selected. The explicit routing is not suitable for large networks.

Constraint routing is a more sophisticated type of explicit routing. Constraint-based routing use a new generation of routing algorithms (e.g. OSPF-TE) taking in account more link parameters than traditional routing algorithms. Unfortunately the existing constraint routing algorithms route only single flows and does not consider the other flows routing. The constraint routing can be affected by bandwidth allocations as well.

2. Bandwidth Constraints Models

There are three bandwidth constraints models proposed up today:

- Maximum Allocation Model (MAM) [3] - the maximum allowable bandwidth usage of each class type (CT) is explicitly specified,
- Russian Doll Model (RDM) [4] - the maximum allowable bandwidth usage is done cumulatively by grouping successive CTs according to priority

to Contents, page 1

classes,

- Maximum Allocation with Reservation Model (MAR) [5] - it is similar to MAM in that a maximum bandwidth allocation is given to each CT. However, through the use of bandwidth reservation and protection mechanisms, CTs are allowed to exceed their bandwidth allocations under conditions of no congestion but revert to their allocated bandwidths when overload and congestion occurs.

Although the comparison of different routing technique in MPLS has been investigated in [6] the bandwidth allocation model has not be considered.

2.1. MAM Model

Maximum Allocation Bandwidth Constraints Model is defined in the following manner:

Assume that Maximum Number of Bandwidth Constraints Θ is equal to Maximum Number of Class-Types Ψ:

$$\Theta = \Psi = 8\,. \tag{1}$$

For each value of n in the range $0 \leq n \leq (\Psi - 1)$:

$$B_{RCT_n} \leq BC_n \leq B_{MaxR}\,, \tag{2}$$

and

$$\sum_{n=0}^{\Psi-1} \mathrm{B}_{\mathrm{RCT_n}} \leq \mathrm{B}_{\mathrm{MaxR}}\,, \tag{3}$$

where B_{RCTn} is the Bandwidth Reserved for Class Type n, BC_n is the Bandwidth Constraint for Class Type n a B_{MaxR} is the Maximum Reservable Bandwidth. The sum of Bandwidth Constraints theoretically may exceed the B_{MaxR}, so that the following relationship may hold true:

$$\sum_{n=0}^{\Psi-1} BC_n > B_{MaxR}\,, \tag{4}$$

but usually the sum of Bandwidth Constraints will be equal to (or below) the Maximum Reservable Bandwidth

$$\sum_{n=0}^{\Psi-1} BC_n \leq B_{MaxR}\,, \tag{5}$$

2.2. RDM Model

RDM model is defined in a similar manner. We expect condition (1).

Then for each value of b in the range $0 \leq b \leq (\Psi$ - 1):

$$\sum_{n=b}^{\Psi-1} \mathrm{B}_{\mathrm{RCT_n}} \leq BC_b\,. \tag{6}$$

Let:

$$BC_0 = B_{MaxR}\,, \tag{7}$$

then

$$\sum_{n=0}^{\Psi-1} \mathrm{B}_{\mathrm{RCT_n}} \leq B_{MaxR}\,. \tag{8}$$

3. Simulation Experiments

To investigate a performance of constraint based routing in MPLS from the point of QoS provisioning for QoS sensitive telecommunication services and efficiency of bandwidth utilization the special simulation model was proposed. All simulations have been done using ns-2.

3.1. Simulation Model

All simulations use the same network architecture. A simulation network topology is shown in Fig. 1. All traffic flows are generated by Source node (node 0) and they are routed to four destination nodes (nodes 8 – 11). The MPLS network is composed of an Ingress Label Edge Router (I-LER), an Egress Label Edge Routers (E-LER) and five Label Switching Routers (LSRs). The MPLS routers are interconnected with 2 Mbps or 10 Mbps links. The external link between Source node (node 0) and I-LER (node 1) is 100 Mbps, the links between E-LER (node 7) and destination nodes (nodes 8 – 11) are 10 Mbps. Ingress and egress Label Edge Routers (node 1 and node 7) are interconnected by two different network segments. The shorter upper segment contains only two Label Switch Routers (LSR2 and LSR3) and it is preferred by classical hop-by-hop routing algorithms. Theoretically, in non congested situation, the packets routed through the upper segment have a smaller delay; therefore the upper segment should be preferred by real time services (e.g. voice). The upper segment throughput is 2 Mbps.

The longer bottom segment contains three LSRs (LSR4, LSR5 and LSR6) and one additional line leading to higher delay and therefore it is less suitable for real time services. The bottom segment throughput is 2 Mbps.

The LSRs in both segments are interconnected by intermediate 10 Mbps links that can be used to create an alternative path. This path has a throughput up to 10 Mbps.

As queue management algorithms a CBQ is applied on MPLS nodes and a drop-tail on external nodes.

to Contents, page 1

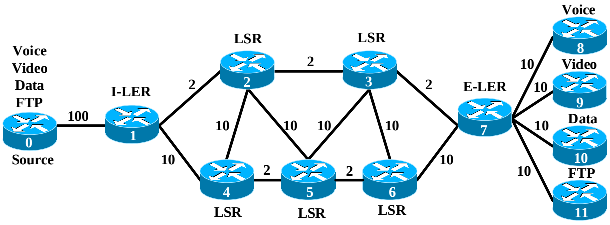


Fig. 1: Network topology (bit rates are in Mbps).

Tab.1: Traffic flows parameters.

Traffic	Bit Rate (kbps)	Packet Size (Bytes)	Traffic Type	Transport Protocol
Voice	952	238	CBR	RTP
Video	1200	250	CBR	RTP
Data	1800	1000	EXP	UDP
FTP	Unlimited	1500	TCP	FTP

There are four traffic flows generated in each simulation representing voice, video, raw data and FTP traffic. The traffic parameters are shown in Tab. 1.

3.2. Hop-by-Hop Routing

The hop-by-hop routing does not provide adequate QoS guaranty. All traffic flows are routed by hop-by-hop routing (shortest path selected). Even there is an available bandwidth on the network, it is not used. All traffic flows compete to 2 Mbps link LER1-LSR2. The link is congested and many packets from all flows are discarded (see Tab. 2).

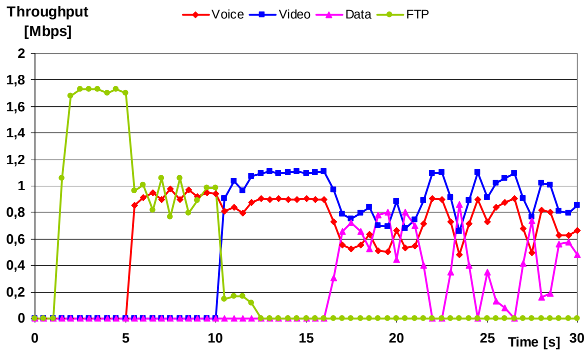
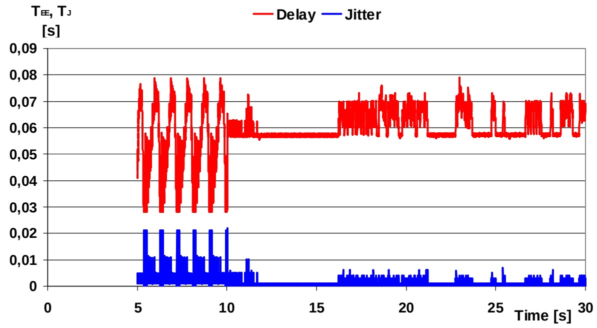
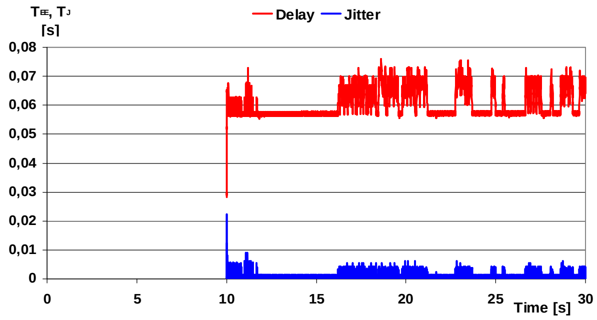


Fig. 2: Throughput of voice, video, data and FTP flows (hop-by-hop routing, WS=10).

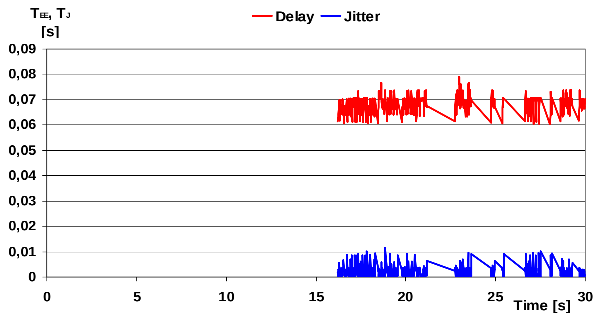
Figure 2 shows the throughput of voice, video, FTP and data flows. The FTP flow starts in time 1 second and up to 5 seconds uses the whole bandwidth. The voice flow starts in time 5 seconds and competes with FTP flow for bandwidth. Because the voice flow uses an UDP and the FTP flow uses a TCP the FTP flow slow down to the rest available bandwidth. The video flow starts in time 10 seconds. Because the voice and video flows require more bandwidth then the link capacity there is no available bandwidth for the FTP flow and the FTP flow slowdown to zero. As the last one starts the data flow in time 15 seconds. As can be seen all flows are affected by congestion and no flow got the required bandwidth. Figure 3 shows delays and jitters of voice, video, data and FTP flows. Because the traffic is not classified by QoS classes all flows have approximately the same delay and jitter (see Tab. 3).



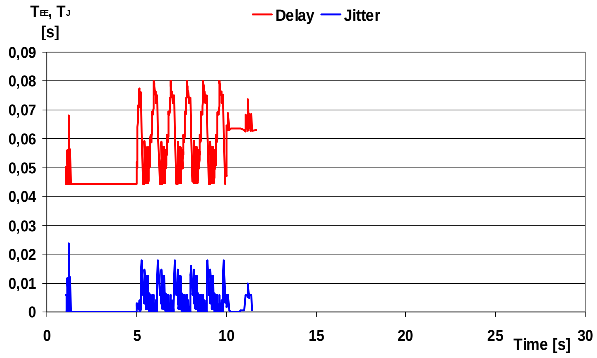
a) Voice



b) Video



c) Data



d) FTP

Fig. 3: Delay and jitter of voice, video, data and FTP flows (hop-by-hop routing, WS=10).

to Contents, page 1

Tab.2: Packet loss (Hop-by-Hop Routing).

Traffic Flow	Packets		
	Sent	Dropped / Retransmissions*	
Voice	12501	2260	18,1 %
Video	12001	2590	21,6 %
Data	1713	951	55,5 %
FTP	984	7*	0,7 %*

Tab.3: Delay and jitter (Hop-by-Hop Routing).

Traffic Flow	Delay [ms]		Jitter [ms]	
	Mean	Max	Mean	Max
Voice	58,3	79,1	0,86	21,95
Video	59,74	76,60	0,76	22,28
Data	68,64	79,00	1,97	11,38
FTP	55,81	80,01	1,30	23,88

3.3. Explicit Routing

The explicit routing allows optimal flow distribution. In proposed example the voice flow is optimally routed via upper segment (shortest delay guaranteed), the video flow via bottom segment (with the second shortest delay) and data and FTP flows are routed via middle segment (the longest delay) marked as: I-LER_LSR4_LSR2_LSR5_ LSR3_LSR6_E-LER (noted as 1_4_2_5_3_6_7 or 1425367). This path selection allows to reach required bandwidth, delay and jitter for all flows. Hence the same routing can be made by constraint routing (see Fig. 4) the values of throughput, delay and jitter shown in Fig. 5 and Fig. 6 and values of packet loss, delay and jitter in Tab. 4 and Tab. 5 are valid for explicit routing as well.

3.4. Constraint Routing

Constraint routing uses routing algorithms taking in account more link parameters than traditional routing algorithms. In optimal situation the constraint routing can select the same routes as optimal explicit routing.

The performance of constraint routing was investigated from the point of:

- optimal path selection,
- flow priority routing,
- bandwidth constraint allocation.

1) Optimal Path Selection

The optimal routing from the point of bandwidth allocation and delay is shown in Fig. 4. The voice, as the high sensitive service uses the shortest path (1_2_3_7), the video the longer path (1_4_5_6_7) and the data and FTP the longest path (1_4_2_5_3_6_7). This proposed path selection prevents link congestion, quarantine required bandwidth for all flows (zero packet loss) and shorter delay for time sensitive services (see Tab. 4 and Tab. 5).

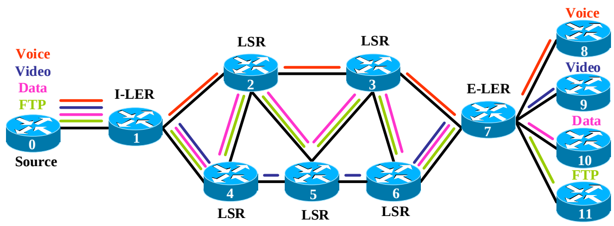


Fig. 4: Optimal traffic routing (constraint routing).

Figure 5 shows a throughput of voice, video, data and FTP flows with RDM constraints model implemented and FTP window size (WS) set to 10. In this case all flows have enough bandwidth, therefore the courses of constant bit rate voice and video flows are flat. The throughput of the data flow is vibrant, while it is generated as a variable bit rate flow with exponential packet distribution. The throughput of the FTP flow is oscillating due to TCP algorithm.

The Fig. 6 shows delay and jitter for a) voice b) video c) data and d) FTP flows. Figure 7 shows the number of transmitted FTP packets vs. FTP window size.

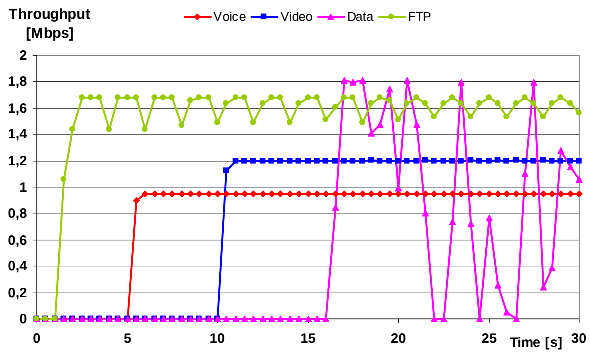


Fig. 5: Throughput of voice, video, data and FTP flows (constraint routing, RDM, WS=10).

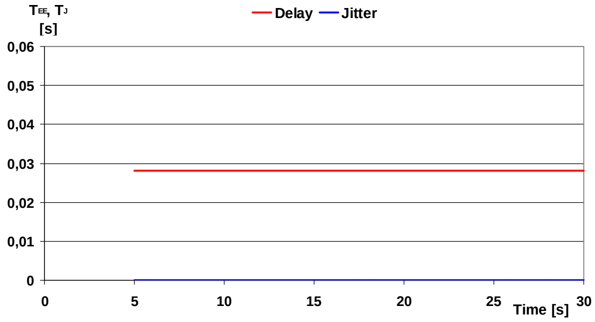
Tab.4: Packet loss (constraint routing, RDM, WS=10).

Traffic Flow	Packets		
	Sent	Dropped / Retransmissions*	
Voice	12501	0	0,0 %
Video	12001	0	0,0 %
Data	1713	0	0,0 %
FTP	3895	0*	0,0 %

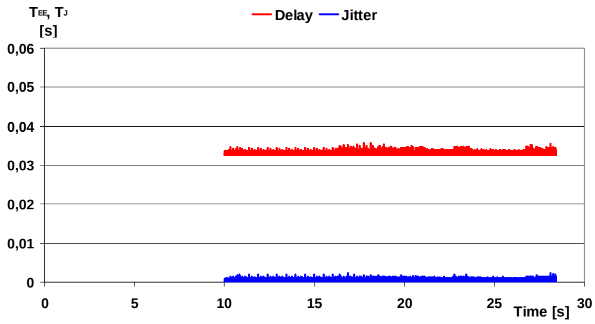
Tab.5: Delay and jitter (constraint routing, RDM, WS=10).

Traffic Flow	Delay [ms]		Jitter [ms]	
	Mean	Max	Mean	Max
Voice	28,06	28,07	0,00	0,006
Video	32,94	35,73	0,32	2,41
Data	46,46	50,35	0,66	3,44
FTP	48,57	53,20	0,06	4,68

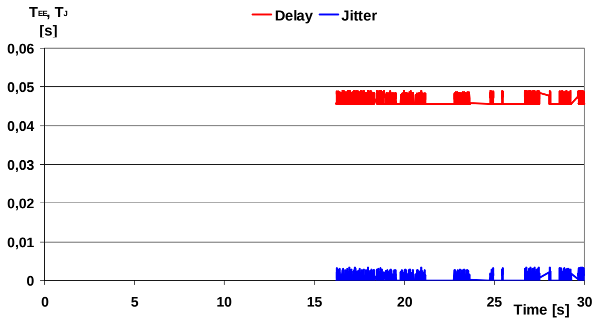
to Contents, page 1



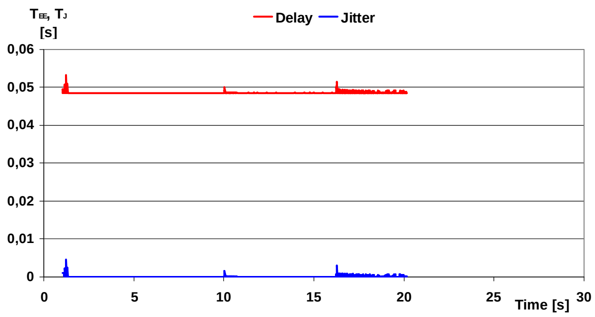
a) Voice



b) Video



c) Data



d) FTP

Fig. 6: Delay and jitter of voice, video, data and FTP flows (constraint routing, RDM, WS=10).

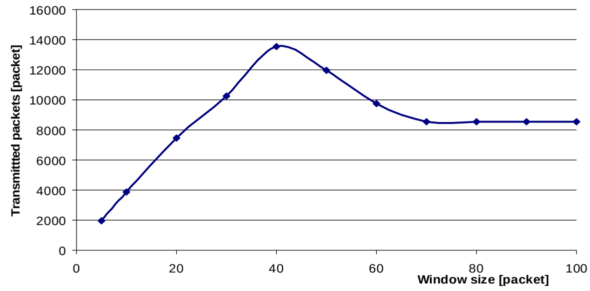


Fig. 7: Number of transmitted packets vs. windows size - FTP flow (constraint routing, RDM).

2) Flow Routing Priority

The results presented above have been achieved for optimal flow order routing i.e. Voice – Video - Data - FTP (Note: the flows start in different order). To show how the routing priority/order and constraint parameters impact the routing, the following simulations have been done.

The first simulation shows the selected path, packet loss, delay and jitter for reservation order Data – FTP – Voice - Video (Tab. 6). Although the low delay sensitive flows (data and FTP) are routed before a high delay sensitive flows (voice and video), due the setting of RDM and setting of bandwidth constraint for Data and FTP flows reasonably high (3 Mbps for FTP), the constraint routing select the same optimal routes as in previous experiment (see Fig. 8).

Figure 9 shows a throughput of voice, video, data and FTP flows with RDM constraints model implemented and FTP window size (WS) set to 40. The Fig. 10 shows delay and jitter for a) voice b) video c) data and d) FTP flows

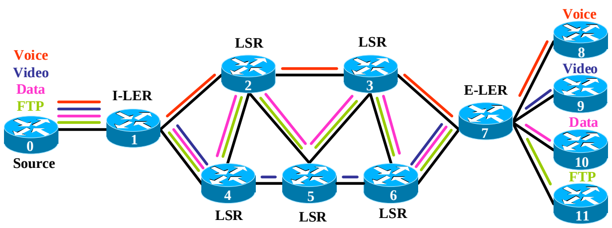


Fig. 8: Traffic routing (constraint routing, constraint bandwidth FTP = 3 Mbps, Data-FTP-Voice-Video, RDM, WS=40).

Tab.6: Routing order Data-FTP-Voice-Video (constraint routing, constraint bandwidth for FTP = 3000 kbps, RDM, WS=40).

	Voice	Video	Data	FTP
Reservation order	3	4	1	2
Req. bandwidth [kbps]	952	1200	1800	Not lim.
Constraint bandw. [kbps]	1000	1200	1800	3000
Selected path [nodes]	1237	14567	1425367	1425367

to Contents, page 1

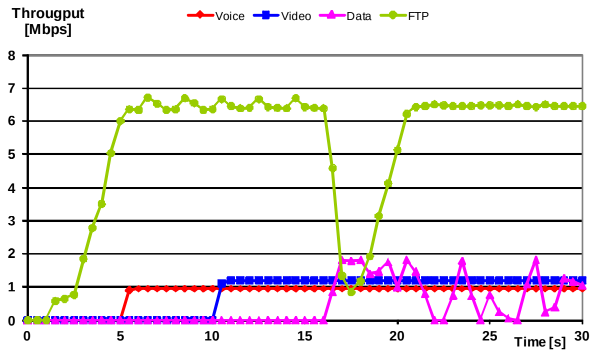


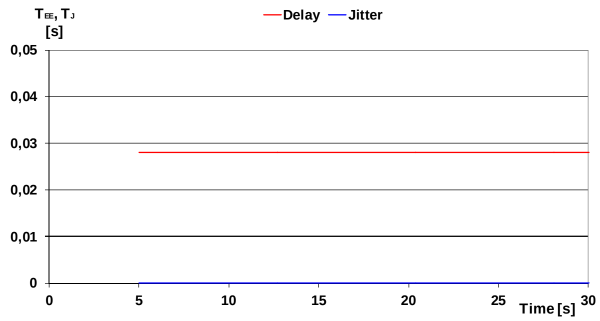
Fig. 9: Throughput of voice, video, data and FTP flows (constraint routing, constraint bandwidth FTP = 3 Mbps, Data-FTP-Voice-Video, RDM, WS=40).

Tab.7: Packet loss (constraint routing, constraint bandwidth FTP = 3 Mbps, Data-FTP-Voice-Video, RDM, WS=40).

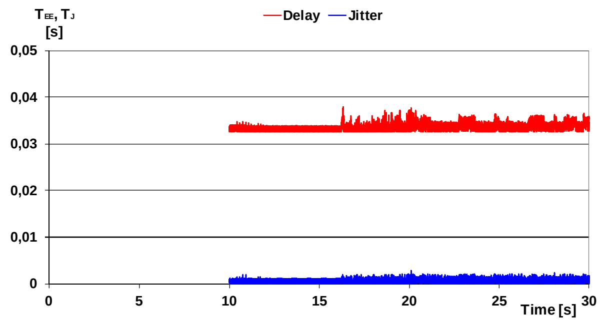
Traffic Flow	Packets		
	Sent	Dropped / Retransmissions*	
Voice	12501	0	0,0 %
Video	12001	0	0,0 %
Data	1713	0	0,0 %
FTP	13189	1*	0,0 %*

Tab.8: Delay and jitter (constraint routing, constraint bandwidth FTP = 3 Mbps, Data-FTP-Voice-Video, RDM, WS=40).

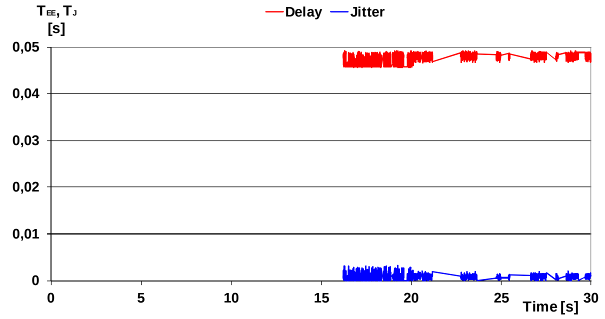
Traffic Flow	Delay [ms]		Jitter [ms]	
	Mean	Max	Mean	Max
Voice	28,06	28,07	0,000	0,006
Video	33,61	37,87	0,482	2,786
Data	47,43	49,09	0,694	3,245
FTP	48,73	55,60	0,126	6,521



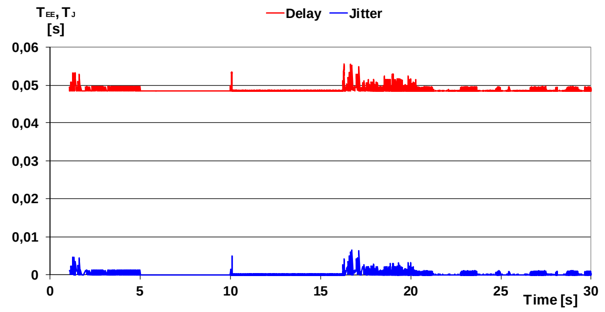
a) Voice



b) Video



c) Data



d) FTP

Fig. 10: Delay and jitter of voice, video, data and FTP flows (constraint routing, constraint bandwidth FTP = 3 Mbps, Data-FTP-Voice-Video, RDM, WS=40).

In the next simulation the bandwidth constraint for FTP flow was decreased to 1 Mbps. This lead to situation that the FTP flow was routed through the shortest path 1_2_3_7, the delay sensitive voice flow through the longer path 1_4_5_6_7 and the video and data flows together via the longest path 1_4_2_5_3_6_7 (see Fig. 11 and Tab. 9).

Although the different route selection does not lead to packet dropping (see Tab. 10) it significantly decrease the throughput of TCP flow (from 13189 packets to 3671 packets – see Fig. 9 and Fig. 12 or/and Tab. 7 and Tab. 10) and increase the delay of voice and video flows. The throughputs of voice, video, data and FTP flows are shown in Fig. 12. The delays and jitters of a) voice b) video c) data and d) FTP flows are shown in Fig. 13.

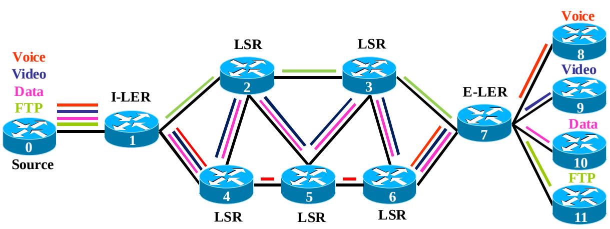


Fig. 11: Traffic routing (constraint routing, constraint bandwidth FTP = 1 Mbps, Data-FTP-Voice-Video, RDM, WS=40).

to Contents, page 1

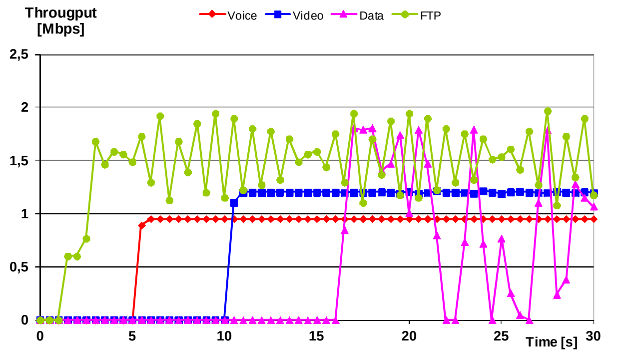


Fig. 12: Throughput of voice, video, data and FTP flows (constraint routing, constraint bandwidth FTP = 1 Mbps, Data-FTP-Voice-Video, RDM, WS=40).

Tab.9: Routing order Data-FTP-Voice-Video (constraint routing, constraint bandwidth FTP = 1 Mbps, RDM, WS=40).

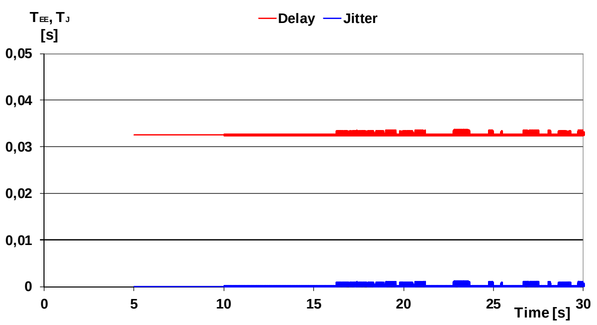
Parameter	Voice	Video	Data	FTP
Rezervation order	3	4	1	2
Req. bandwidth [kbps]	952	1200	1800	Not lim.
Constraint bandw. [kbps]	1000	1200	1800	1000
Selected path [nodes]	14567	1425367	1425367	1237

Tab.10: Packet loss (constraint routing, constraint bandwidth FTP = 1 Mbps, Data-FTP-Voice-Video, RDM, WS=40).

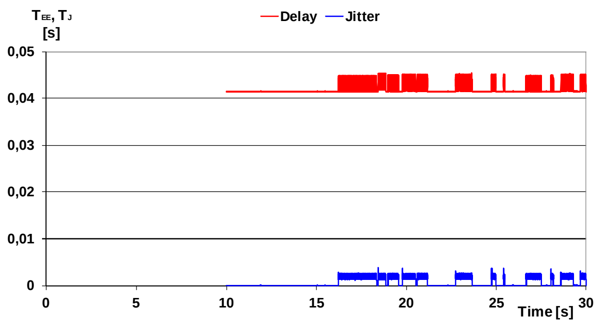
Traffic Flow	Packets		
	Sent	Dropped / Retransmissions*	
Voice	12501	0	0,0 %
Video	12001	0	0,0 %
Data	1713	0	0,0 %
FTP	3671	0*	0,0 %*

Tab.11: Delay and jitter (constraint routing, constraint bandwidth FTP = 1 Mbps, Data-FTP-Voice-Video, RDM, WS=40).

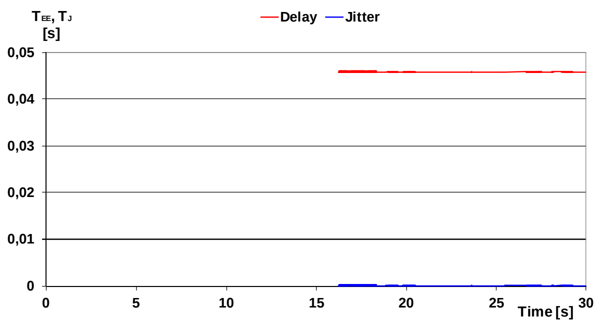
Traffic Flow	Delay [ms]		Jitter [ms]	
	Mean	Max	Mean	Max
Voice	32,59	33,61	0,185	1,122
Video	42,13	45,37	0,667	3,780
Data	45,70	45,97	0,042	0,291
FTP	54,19	68,32	1,298	24,000



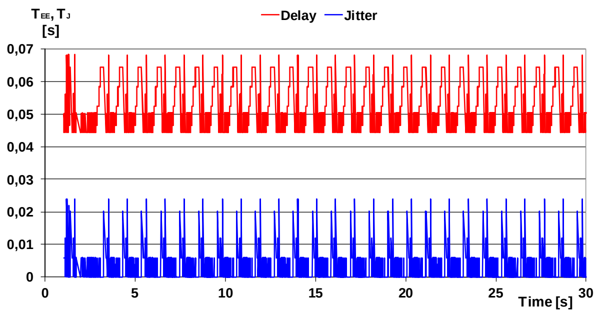
a) Voice



b) Video



c) Data



d) FTP

Fig. 13: Delay and jitter of voice, video, data and FTP flows (constraint routing, constraint bandwidth, FTP = 1 Mbps, Data-FTP-Voice-Video, RDM, WS=40).

In the last simulation the order and constraints were changed. The flows were routed in order FTP-Data-Voice-Video Due to the bandwidth constraints for FTP and Data flows were set only to 1,5 Mbps (see Tab. 12). The FTP flow was routed via the upper shortest path 1_2_3_7, data flow via the second shortest bottom path and delay sensitive voice and video via the longest path 1_4_2_5_3_6_7 (see Fig. 14). The throughputs of voice, video, data and FTP flows are shown in Fig. 15. The delays and jitters of a) voice b) video c) data and d) FTP flows are shown in Fig. 16 and Tab. 14.

to Contents, page 1

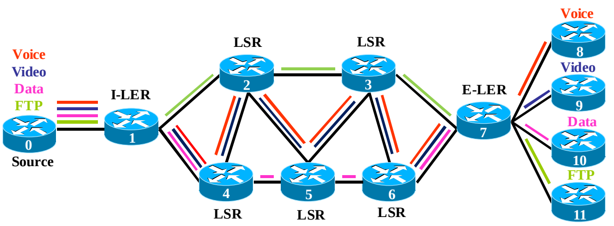


Fig. 14: Traffic routing (constraint routing, constraint bandwidth FTP = 1,5 Mbps, FTP-Data-Voice-Video, RDM, WS=40).

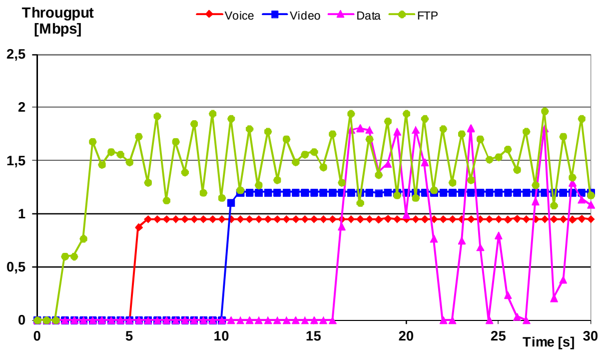


Fig. 15: Throughput of voice, video, data and FTP flows (constraint routing, constraint bandwidth, FTP = 1,5 Mbps, FTP-Data-Voice-Video, RDM, WS=40).

Tab.12: Routing order FTP-Data-Voice-Video (constraint routing, constraint bandwidth Data=1,5 Mbps, FTP = 1,5 Mbps, RDM, WS=40).

Parameter	Voice	Video	Data	FTP
Rezervation order	3	4	2	1
Req. bandwidth [kbps]	952	1200	1800	Not lim.
Constraint bandw. [kbps]	1000	1200	1500	1500
Selected path [nodes]	1425367	1425367	14567	1237

Tab.13: Packet loss (constraint routing, constraint bandwidth FTP = 1,5 Mbps, FTP-Data-Voice-Video, RDM, WS=40).

Traffic Flow	Packets		
	Sent	Dropped / Retransmissions*	
Voice	12501	0	0,0 %
Video	12001	0	0,0 %
Data	1713	0	0,0 %
FTP	3671	1*	0,0 %*

Tab.14: Delay and jitter (constraint routing, constraint bandwidth FTP = 1,5 Mbps, FTP-Data-Voice-Video, RDM, WS=40).

Traffic Flow	Delay [ms]		Jitter [ms]	
	Mean	Max	Mean	Max
Voice	41,44	42,38	0,105	1,029
Video	41,48	42,25	0,109	0,837
Data	40,52	40,92	0,071	0,442
FTP	54,17	68,32	1,298	24,000

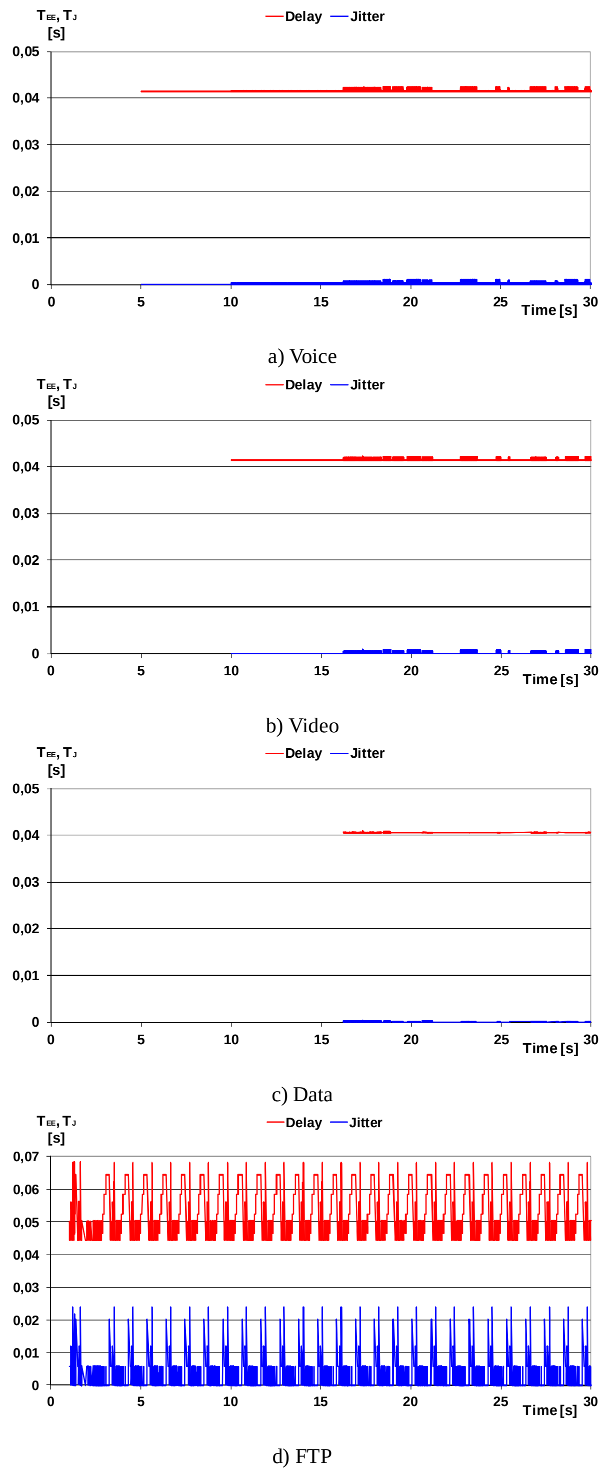


Fig. 16: Delay and jitter of voice, video, data and FTP flows (constraint routing, constraint bandwidth FTP = 1,5 Mbps, FTP-Data-Voice-Video, RDM, WS=40).

As can be seen from Fig. 16 and Tab. 14 the routing of voice traffic over the longest path lead to highest end-to-end delay and jitter. Even in this experiment delay and jitter for voice (and video) traffic are acceptable, the non optimal routing can in some situations lead to non acceptable high values of delay or jitter.

to Contents, page 1

3) Bandwidth Constraints Models

The all previous simulations used a RDM bandwidth constraint model. RDM allows reuse of non allocated bandwidth from higher priority class types by traffic flow with lower priority class type. The disadvantage of RDM is a necessity to maintain preemption in all network nodes. MAM explicitly specify maximum allowable bandwidth usage of each class type and the bandwidth can not be shared among the different class types. This should be considered during network design process.

Table 15 shows the MPLS path allocation proposed by constraint routing algorithm for different bandwidth allocations for specified class type (CT). As can be seen, if not enough bandwidth is allocated to the CT on the link, the constraint routing algorithm chose the less preferable link. If the process fails on all links, the traffic flow is routed by hop-by-hop routing. This can lead to congestion on most preferable link usually used by the most time critical services.

Tab.15: Path selection (CR, MAM, WS=40).

CT Bandw. Allocation [%]	Voice	Video	Data	FTP
85	1237	14567	1425367	1425367
75	1237	14567	1425367	1425367
65	1237	14567	1425367	1425367
55	1237	1425367	1425367	Hop-by-hop
45	1425367	1425367	1425367	Hop-by-hop
35	1425367	1425367	Hop-by-hop	Hop-by-hop
25	1425367	1425367	Hop-by-hop	Hop-by-hop
15	1425367	Hop-by-hop	Hop-by-hop	Hop-by-hop

Figure 17 to Fig. 19 show the throughput of voice, video, data and FTP flows when the 85 %, 55 % or 35 % of the bandwidth is allocated to the CT used by traffic flows. As can be seen in Fig. 18, when the bandwidth constraint is set to 55 %, the FTP flow shares the 2 Mbps link with the voice flow (see Tab. 15) that degrades the QoS parameters of both of them. The number of loss packet is depicted in Tab. 16 to Tab. 18.

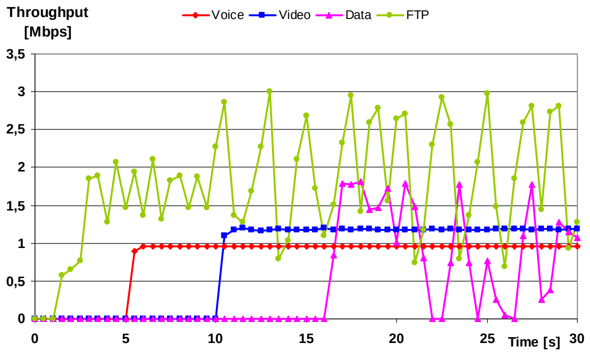


Fig. 17: Throughput of voice, video, data and FTP flows (constraint routing, MAM 85 %, WS=40).

Tab.16: Packet loss (CR, MAM 85 %, WS=40).

Traffic Flow	Packets		
	Sent	Dropped / Retransmissions*	
Voice	12501	0	0,0 %
Video	12001	191	1,6 %
Data	1713	0	0,0 %
FTP	4496	5*	0,1 %

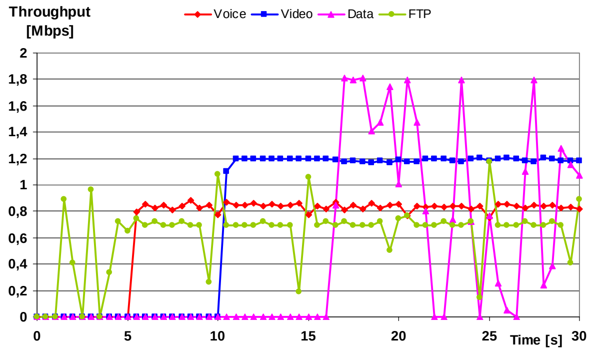


Fig. 18: Throughput of voice, video, data and FTP flows (constraint routing, MAM 55 %, WS=40).

Tab.17: Packet loss (CR, MAM 55 %, WS=40).

Traffic Flow	Packets		
	Sent	Dropped / Retransmissions*	
Voice	12501	1531	12,2 %
Video	12001	91	0,7 %
Data	1713	0	0,0 %
FTP	1675	3*	0,2 %

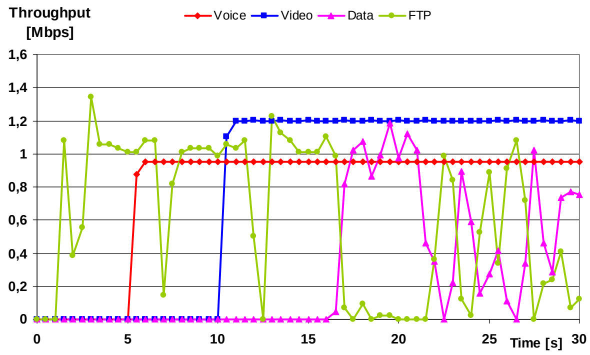


Fig. 19: Throughput of voice, video, data and FTP flows (constraint routing, MAM 35 %, WS=40).

Tab.18: Packet loss (CR, MAM 35 %, WS=40).

Traffic Flow	Packets		
	Sent	Dropped / Retransmissions*	
Voice	12501	0	0,0 %
Video	12001	0	0,0 %
Data	1682	618	36,7 %
FTP	1667	10*	0,6 %

to Contents, page 1

4. Conclusion

The proposed experiments accomplished that constraints routing can be used as traffic engineering and QoS provisioning tool allowing providing QoS parameters for QoS sensitive services.

Even its good performance, it has been shown that in some situations such as non optimal order of the routed flows selected or inadequate bandwidth allocation with MAM bandwidth constraints model implemented the routing can fail or non optimal routes leading to QoS parameter degradation can be proposed.

Constraint-based routing should be implemented in conjunction with other QoS provisioning methods such as traffic flow differentiation according QoS requirements, different processing in network nodes (different Per-Hop Behaviors) or bandwidth reservation for different service classes.

Acknowledgements

This paper has been supported by the projects VEGA No. 1/0565/09 and ITMS 26240120029.

References

[1] ROSEN, E; VISWANATHAN, A; CALLOLN, A. *Multiprotocol Label Switching Architecture*, RFC 3031, IETF, January 2001. pp. 1-61

[2] HALAS, Michal; KYRBASHOV Bakyt; VOZNAK Miroslav. Factors influencing voice quality in VoIP technology. In *9th International Conference on Informatics'2007*, 21-22. June 2007, Bratislava, pp. 32-35. ISBN 978-80-969243-7-0.

[3] Le FAUCHEUR, F.; LAI, W. *Maximum Allocation Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering*, RFC 4125, IETF, June 2005. pp. 1-10

[4] Le FAUCHEUR, F. *Russian Dolls Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering*, RFC4127, IETF, June 2005. pp. 1-12

[5] ASH, J. *Max Allocation with Reservation Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering & Performance Comparisons*, RFC 4126, IETF, June 2005. pp. 1-20

[6] MEDVECKY, Martin. The Impact of Traffic Routing in MPLS Networks to QoS. In *RTT 2008 - 9th International Conference Research in Telecommunication Technologies*, Vyhne, Slovak Republic, 10.9.-12.9.2008, 2008. pp. 187-191. ISBN 978-80-227-2939-0.

About Author

Martin MEDVECKY was born in Trencianske Bohuslavice, Slovakia. He received his M.Sc. in "Telecommunications Technology – Telecommunications Systems" in 1991, Ph.D. in "Telecommunications" in 1998 and Assoc Prof. in "Telecommunications" in 2008 from Slovak University of Technology in Bratislava, Slovakia. His research interests include broadband switching, QoS in broadband networks and telecommunication network management.

to Contents, page 1

TREE-ORIENTED DATA ACQUISITION ARCHITECTURE

Vit NOVOTNY[1]

[1]Department of Telecommunication, Faculty of Electrical Engineering and Communication, Brno University of Technology, Purkynova 118, 612 00 Brno, Czech Republic

novotnyv@feec.vutbr.cz

Abstract. *Hierarchical system of network nodes is suitable solution how to collect and how to pre-process data from large amount of end-nodes. By contrast to flat (one layer) architecture there are special intermediary nodes used and they are called summarization nodes. These special nodes have to be suitably placed in the network to enable efficient data collection and their number in the hierarchy is one of the key parameters of the architecture. The article deals with the tree architecture design, with its optimisation and with the problem of limited number of summarization nodes.*

Keywords

Data acquisition, tree architecture, multicast, sensor networks, SSM, multimedia, RTP/ RTCP.

1. Introduction

There are several ways of data collection and processing in the network environment. First model is called centralized, where there is only one data centre where all pieces of information are collected, processed and available. The second one is called hierarchical system created by a tree of servers with local data and using links among them the requested information can be found through the tree hierarchy. The third model is distributed model where the pieces of information are distributed among equivalent data centres and using one sophisticated directory services with one account requested information can be obtained.

Many applications are using plain centralized model. There is no problem with data acquisition and with data processing provided the number of data sources is fairly low and data flows are weak and low frequent. When these conditions are not fulfilled either the centre itself or data links to the centre can be overloaded or allowed data transmission frequency is very low. When the data acquisition is auxiliary procedure of the service, the available bandwidth for such procedure is strictly limited and the situation becomes even worse. This is the case of applications like IP-TV where the main procedure of the service is the multimedia streaming using RTP protocol and the multicast transmission and the session quality parameter collection using RTCP protocol is an optional though useful supplementary service [1], [2], [3]. The transmission capacity of RTCP is limited for 5 % of total service bandwidth and it causes large delays in sending RTCP (feedback) data from each receiver for large-scale media streaming services based on Source-Specific Multicast (SSM), [7]. Similar problem arises also with other applications focused on data acquisition in the case of large-scale systems.

2. Hierarchical Data Acquisition System

To combat the problem the hierarchical system for data acquisition has been proposed in [5], [6] and modified in [7]. In addition to the data centre and data sources such tree contains special nodes called summarization nodes, see Fig. 1.

The data is periodically sent from data sources (terminals or sensors) to assigned summarization node. The summarization node aggregates data from a group of terminals of the size n_B and again periodically sends to assigned summarization node at the higher level. The summarization nodes are also organized into groups of size n_S. Structure of Receiver Summary Information (RSI) message was specified in [5]. The message includes sub-report blocks (SRB) that contain distribution information about particular features like a packet loss or a jitter.

To enable efficient transmission of information about the session from the data centre to the terminals an extension of original RTCP specification in the form of Extend Report (XR) message had to be adopted [3]. The

to Contents, page 1

XR RTCP summarization packet consists basic information for the terminals mainly how to calculate the message transmission period. In the case of SSM (Source-Specific Multicast) service the message is sent in multicast manner so that together with the summarization method it decreases the overhead and saves the bandwidth.

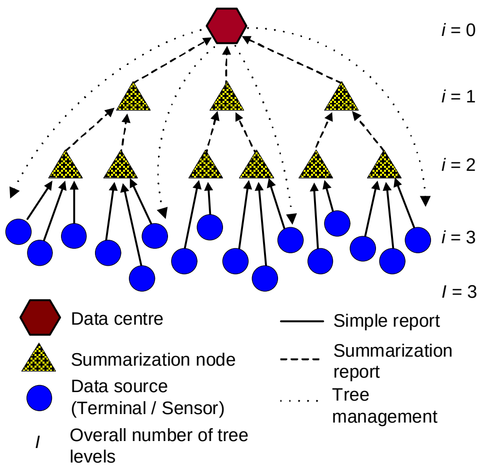


Fig. 1: Tree design for large-scale data acquisition.

Optionally in addition to the summarization process (when detailed information is lost) the summarization nodes can store detailed information obtained from terminals or lower level summarization nodes for some time period to allow the data centre to get detailed information about particular terminal or group of terminals when necessary.

Because of large group of terminals division into a big number of smaller groups the bandwidth restriction is not the problem and the message transmission period of the terminals remains fairly low even if the overall number of terminals rises. Especially this is the case of multimedia multicast sessions which can vary substantially in size. The overall delay that is bounded by the time instant when data is generated (or measured) in the terminal (sensor) and by the time instant when the data is received in the data centre consists of particular transmission delays between transmission instants of adjacent layers in the tree. The situation is depicted in Fig. 2:

When the tree consists of I layers, i.e. (I-1) summarization layers and one terminal layer, a formula for the overall delay T_R between data generation (measurement) and its reception in the data processing centre can be derived:

$$T_{\mathrm{R}} = \tau_{\mathrm{MT}} + \sum_{i=1}^{I-1} \tau_i \,, \tag{1}$$

provided the transport delay through the network is neglected. Variable τ_{MT} is the delay between measurement (data generation) and transmission instants and τ_i is the delay between summarized message transmission instants at linked summarization nodes in adjacent layers.

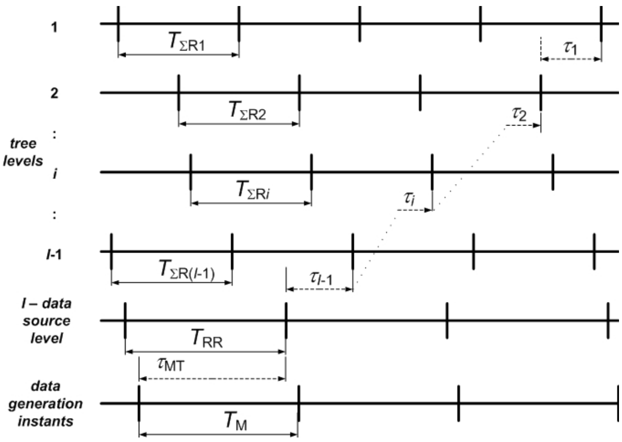


Fig. 2: Time instants of data generation and message transmissions in hierarchical data acquisition system.

The worst case for the delay will be when all summarization nodes at all levels of the tree and also the terminals (sensors) are synchronized, i.e. all of them transmit messages almost at the same time instants. Then the formula (1) will convert to

$$T_{\mathrm{RTW}} = T_{\mathrm{RR}} + \sum_{i=1}^{I-1} T_{\Sigma\mathrm{R}i} \,. \tag{2}$$

Provided the transmission periods are the same through whole tree the formula (2) changes into the form

$$T_{\mathrm{RTW}} = T_{\mathrm{RR}} + \left(I-1\right) T_{\Sigma\mathrm{R}} \,, \tag{3}$$

where T_{RR} is the transmission period of the group of terminals (it depends on the number n_{B} of terminals in the group, message length and the allocated bandwidth, [7]), I is the number of levels in the tree (it depends on the total number of terminals, on the number of terminals in the group and on the number of summarization nodes in the group) and $T_{\Sigma\mathrm{R}}$ is the message transmission period of the summarization node group (it depends on the number n_{S} of summarization nodes in the group, summarization message length and the allocated bandwidth, [7]).

Now several problems come out. First group of problems are how to manage the tree when the number of terminals rises or declines, how to keep it in balanced form and how to minimize the total delay specified by (3) . In addition to this the problem of the number of required summarization nodes should be addressed. At the beginning of our research we considered that the summarization nodes are only terminals with special functionality [7]. It was found that there would be lot of overhead with the management of such tree especially when the tree is variable in a large extent, i.e. the terminals will enter and leave the session frequently; this is the case of multimedia streaming sessions. Also this functionality would require additional power and energy

to Contents, page 1

that is unwanted issue especially in the case of wireless terminals (sensors) with very limited computational power and energy. Therefore in later research ([8], [10]) the summarization nodes are considered as special nodes (or software modules) that are managed by the service provider. Such summarization nodes have higher computational power, larger storage capacity for temporary data and fixed location. This last feature is very important when tree structure is established according to the location of terminals, [10].

3. Tree Optimization

When the service provider intents to implement a service based on the tree architecture described above before implementation some initial conditions have to be considered: bandwidth (or maximum data flow) allocated for the data acquisition BW_A (it will be allocated for each group of terminals or summarization nodes), expected number of data sources (terminals) n_T, maximum period (or delay) of data collection T_{Rmax}, length PL_{RR} of plain messages generated by the terminals and the length $PL_{\Sigma R}$ of summarization packets generated by summarization nodes. Additional constraints can be: maximum overall number of available summarization nodes N_{STmax}, minimum periods of message transmission in a group of terminals T_{RRmin} and in a group of summarization nodes $T_{\Sigma Rmin}$ and some others. The goal is to find such tree that meets all of these conditions and restrictions.

Equation (3) shows how to calculate the largest overall delay T_R (and also the maximum time period of data acquisition) between data generation (measurement) in terminals (data sources) and its reception in the data processing centre. It can be worked out in more detailed form:

$$\begin{aligned} T_{RTW} &= T_{RR} + (I-1)T_{\Sigma R} = \tau_R n_B + (I-1)\tau_\Sigma n_S = \\ &= (BW_A)^{-1}\left[PL_{RR} n_B + (I-1) PL_{\Sigma R} n_S\right], \end{aligned} \tag{4}$$

where I is a number of tree levels, τ_R is a time interval consumed by one message send by a terminal and τ_Σ is a time interval consumed by one summarization message generated by summarization node, n_B is the number of terminals in one group of terminals, n_S is the number of nodes in the group of summarization nodes (the rest of symbols are explained in the text above).

The number of levels with summarization nodes in the tree, i.e. the value (I-1), can be calculated from the condition

$$n_S^{(I-2)} < {}^{n_T}\!/_{n_B} \le n_S^{(I-1)}. \tag{5}$$

Then

$$\log_{n_S}\left({}^{n_T}\!/_{n_B}\right) \le (I-1) < \log_{n_S}\left({}^{n_T}\!/_{n_B}\right) + 1. \tag{6}$$

As I is an integer number the nearest higher integer will be

$$(I-1) = \log_{n_S}\left(\frac{n_T}{n_B}\eta_I\right); \quad \eta_I \in \langle 1, n_S). \tag{7}$$

Then (4) changes into form

$$\begin{aligned} T_{RTW} &= \tau_R n_B + (I-1)\tau_\Sigma n_S = \\ &= \tau_R n_B + (\tau_\Sigma n_S)\left\lceil \log_{n_S}\left(\frac{n_T}{n_B}\eta_I\right)\right\rceil. \end{aligned} \tag{8}$$

so that the optimization process of delay minimization consists of seeking the numbers n_B and n_S. An example of the delay course according to (8) is shown in Fig. 3:

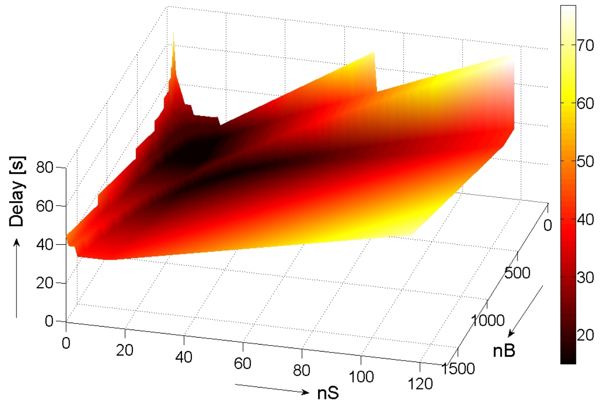


Fig. 3: Course of the worst-case total delay according to (8).

When discontinuous function (8) is replaced by continuous one (without correction parameter η_I), we obtain expression

$$T_{RTW} = \tau_R n_B + (\tau_\Sigma n_S)\left\lceil \log_{n_S}\left(\frac{n_T}{n_B}\right)\right\rceil. \tag{9}$$

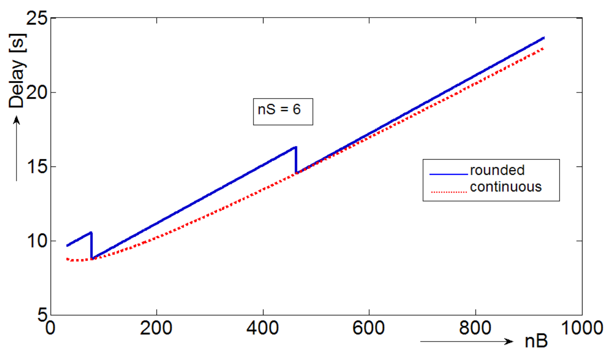


Fig. 4: Comparison of discontinuous (8) and continuous (9) representations of the worst-case delay.

The worst-case total delay values obtained from the optimization process with continuous function are quite close to and always a better than when

to Contents, page 1

discontinuous function is considered (due to the fact that $\eta_I \geq 1$), see Fig. 4:

The goal of optimization is to find its global extreme (minimum) in this region. Global extreme can be located either in local extremes of the function or at the boundary of definition domain. The function is continuous in whole region and smooth, therefore first and also second derivatives can be calculated and stationary points of the function can be found:

$$\frac{\partial T_{\mathrm{RTW}}}{\partial n_{\mathrm{B}}} = \tau_{\mathrm{R}} - \tau_{\Sigma} \frac{n_{\mathrm{S}}}{n_{\mathrm{B}} \ln n_{\mathrm{S}}}, \tag{10}$$

$$\frac{\partial T_{\mathrm{RTW}}}{\partial n_{\mathrm{S}}} = \tau_{\Sigma} \frac{\ln\left(\frac{n_{\mathrm{T}}}{n_{\mathrm{B}}}\right)}{\ln^2 n_{\mathrm{S}}} \left(\ln n_{\mathrm{S}} - 1\right). \tag{11}$$

Stationary points are the candidates for local extremes and they can be calculated from the conditions that first derivatives (10) and (11) are put equal zero and the results are:

$$n_{\mathrm{Ss1}} = \mathrm{e} \text{ (i.e. 2.71828...)} + \eta_{\mathrm{S}} = 3 \tag{12}$$

and when non-rounded $n_{\mathrm{Ss1;}}$ is used for n_{Bs1} calculations

$$n_{\mathrm{Bs1}} = \frac{\tau_{\Sigma}}{\tau_{\mathrm{R}}} \frac{n_{\mathrm{Ss1}}}{\ln n_{\mathrm{Ss1}}} = \frac{\tau_{\Sigma}}{\tau_{\mathrm{R}}} \mathrm{e} + \eta_{\mathrm{B}} = \frac{PL_{\Sigma\mathrm{R}}}{PL_{\mathrm{RR}}} \mathrm{e} + \eta_{\mathrm{B}}, \quad \eta_{\mathrm{B}} \in \langle -0.5; +0.5 \rangle. \tag{13}$$

To prove, whether the local minimum was found, it is necessary to check sufficient conditions for the existence of local minimum

$$D_1 = \left.\frac{\partial^2 T_{\mathrm{RTW}}}{\partial n_{\mathrm{B}}^2}\right|_{\substack{n_{\mathrm{B}}=n_{\mathrm{Bs}} \\ n_{\mathrm{S}}=n_{\mathrm{Ss}}}} \geq 0, \quad D_2 = \left|\begin{matrix} \frac{\partial^2 T_{\mathrm{RTW}}}{\partial n_{\mathrm{B}}^2} & \frac{\partial^2 T_{\mathrm{RTW}}}{\partial n_{\mathrm{B}} \partial n_{\mathrm{S}}} \\ \frac{\partial^2 T_{\mathrm{RTW}}}{\partial n_{\mathrm{B}} \partial n_{\mathrm{S}}} & \frac{\partial^2 T_{\mathrm{RTW}}}{\partial n_{\mathrm{S}}^2} \end{matrix}\right|_{\substack{n_{\mathrm{B}}=n_{\mathrm{Bs}} \\ n_{\mathrm{S}}=n_{\mathrm{Ss}}}} > 0. \tag{14}$$

When the results (12) and (13) are used in (14) we get:

$$D_1 = \frac{\tau_{\mathrm{R}}^2}{\tau_{\Sigma} \mathrm{e}}, \quad D_2 = \frac{\tau_{\mathrm{R}}^2}{\mathrm{e}^2} \left[\ln\left(n_{\mathrm{T}} \frac{\tau_{\mathrm{R}}}{\tau_{\Sigma}} \right) - 1 \right]. \tag{15}$$

The inequality $D_1 > 0$ is always met and the inequality D_2 will be fulfilled when

$$n_{\mathrm{T}} \frac{\tau_{\mathrm{R}}}{\tau_{\Sigma}} > \mathrm{e}. \tag{16}$$

Again in the example of RTCP presented in [7] the length of receiver report PL_{RR} was 736 bits and the length of summarization report $PL_{\Sigma\mathrm{R}}$ was 11296 bits. In the case when the same link bandwidths are assigned both to terminals and summarization nodes (16) has the form

$$n_{\mathrm{T}} \frac{\tau_{\mathrm{R}}}{\tau_{\Sigma}} = n_{\mathrm{T}} \frac{PL_{\mathrm{RR}}}{PL_{\Sigma\mathrm{R}}} = n_{\mathrm{T}} \frac{736}{11296} \approx 0.065 n_{\mathrm{T}} > \mathrm{e}, \quad n_{\mathrm{T}} > 42. \tag{17}$$

This condition is quite easy to meet.

Provided the condition (16) is met the (12) and (13) specify local delay minimum:

$$T_{\mathrm{RTW1}}\left(n_{\mathrm{Bs1}} = \frac{\tau_{\Sigma}}{\tau_{\mathrm{R}}} \mathrm{e} + \eta_{\mathrm{B}}, n_{\mathrm{Ss1}} = 3 \right) = = \tau_{\Sigma} \left[\mathrm{e} + 3 \log_3 \left(\frac{n_{\mathrm{T}}}{\frac{\tau_{\Sigma}}{\tau_{\mathrm{R}}} \mathrm{e} + \eta_{\mathrm{B}}} \right) \right] + \tau_{\mathrm{R}} \eta_{\mathrm{B}}. \tag{18}$$

When minimum periods of message transmission in a group of terminals T_{RRmin} and in a group of summarization nodes $T_{\Sigma\mathrm{Rmin}}$ are required then minimum values n_{Bmin}, n_{Smin} are set:

$$\begin{aligned} n_{\mathrm{Bmin}} &= \frac{BW_{\mathrm{A}}}{PL_{\mathrm{RR}}} T_{\mathrm{RRmin}} + \eta_{\mathrm{Bmin}} = \frac{T_{\mathrm{RRmin}}}{\tau_{\mathrm{R}}} + \eta_{\mathrm{Bmin}}, \\ &\eta_{\mathrm{Bmin}} \in \langle 0,1 \rangle, \\ n_{\mathrm{Smin}} &= \frac{BW_{\mathrm{A}}}{PL_{\Sigma\mathrm{R}}} T_{\Sigma\mathrm{Rmin}} + \eta_{\mathrm{Smin}} = \frac{T_{\Sigma\mathrm{Rmin}}}{\tau_{\Sigma}} + \eta_{\mathrm{Smin}}, \\ &\eta_{\mathrm{Smin}} \in \langle 0,1 \rangle. \end{aligned} \tag{19}$$

Then the absolute minimum will be reached for the smallest $n_{\mathrm{S}} = n_{\mathrm{Smin}}$ and for $n_{\mathrm{B}} = n_{\mathrm{Bmin}}$.

4. Summarization Nodes

The service provider has no unlimited number of summarization nodes available and therefore the overall number of required summarization nodes N_{ST} in the tree hierarchy is also very important parameter and should be optimized. The total number of summarization nodes can be calculated as follows (see Fig. 5:):

$$N_{\mathrm{ST}} = \sum_{i=1}^{I-1} N_{\mathrm{S}i} = \sum_{i=1}^{I-2} n_{\mathrm{S}}^{i} + N_{\mathrm{S}(I-1)} = \frac{n_{\mathrm{S}}^{I-2} - 1}{1 - n_{\mathrm{S}}^{-1}} + N_{\mathrm{S}(I-1)}, \tag{20}$$

where I is the number of levels in the tree, $N_{\mathrm{S}i}$ is the number of summarization nodes at the level i, n_{S} is the number of summarization nodes in one group and $N_{\mathrm{S}(I-1)}$ is the number of summarization nodes at the level I-1.

to Contents, page 1

The parameter n_S is known and therefore the task is to calculate the variable $N_{S(I-1)}$. As shown in Fig. 5:the terminals (sensors) are connected to the summarization nodes at two layers, (I-2) and (I-1) respectively.

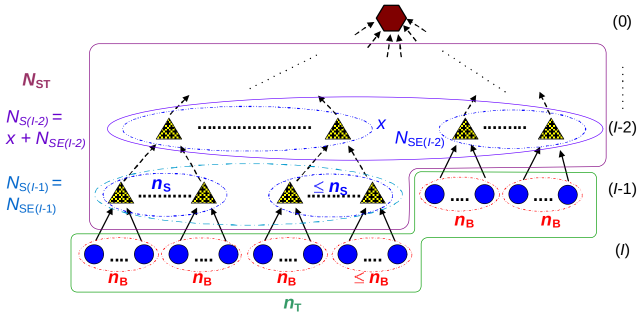


Fig. 5: Tree architecture for data acquisition with highlighting of last three levels.

These summarization nodes can be called summarization endpoint nodes N_{SE} and its total number N_{SET} (T = total) can be expressed by formula

$$N_{SET} = N_{SE(I-2)} + N_{SE(I-1)}, \tag{21}$$

where $N_{SE(I-2)}$ and $N_{SE(I-1)}$ are endpoint summarization nodes at levels (I-2) and (I-1) respectively. As the (I-1) layer is the last layer of the summarization nodes it is clear that $N_{S(I-1)} = N_{SE(I-1)}$. The parameter N_{SET} is an integer figure and it can be calculated by the equation

$$N_{SET} = \frac{n_T}{n_{Bmax}} + \eta_E; \quad \eta_E \in \langle 0,1 \rangle, \tag{22}$$

where n_{Bmax} is the maximum number of terminals in one group.

To obtain the total number of required summarization nodes N_{ST} it is necessary to calculate $N_{SE(I-1)}$. To get this parameter we need $N_{SE(I-2)}$ first. When new terminals are to be added and the current tree is not sufficient, additional layer of summarization nodes has to be added. An appropriate number of summarization nodes x that will loose the terminals for next-layer summarization nodes (where the maximum of summarization nodes therefore can be $x*n_S$) can be calculated from (23):

$$\begin{aligned} N_{SET} &= N_{SE(I-2)} + N_{SE(I-1)} = \\ &= N_{S(I-2)} - x + N_{S(I-1)} = \\ &= n_S^{I-2} - x + N_{S(I-1)} \le n_S^{I-2} - x + x n_S . \end{aligned} \tag{23}$$

The closest larger integer is

$$x = \frac{\frac{n_T}{n_{Bmax}} + \eta_E - n_S^{I-2}}{n_S - 1} + \eta_x; \quad \eta_x \in \langle 0,1 \rangle. \tag{24}$$

Then the total number of required summarization nodes is:

$$\begin{aligned} N_{ST} &= \sum_{i=1}^{I-1} N_{Si} = \sum_{i=1}^{I-2} n_S^i + N_{S(I-1)} = \frac{n_S^{I-2} - 1}{1 - n_S^{-1}} + \\ &+ \frac{N_{SET} - n_S^{I-2}}{1 - n_S^{-1}} + \eta_x = \frac{N_{SET} - 1}{1 - n_S^{-1}} + \eta_x = \\ &= \frac{1}{1 - n_S^{-1}} \left[\frac{n_T}{n_{Bmax}} - 1 \right] + \frac{\eta_E}{1 - n_S^{-1}} + \eta_x . \end{aligned} \tag{25}$$

The formula (25) shows that the overall number of summarization nodes is mainly influenced by total number of receivers n_T and by the number of terminals in one group n_{Bmax}. Parameter n_S does not have big impact on the N_{ST} when $n_s >> 1$.

The rest of the formula (25), i.e. expression $\frac{\eta_E}{1 - n_S^{-1}} + \eta_x$ will be always < 3 (even when $n_S = 2$, which is the smallest value, and both $\eta_x \to 1$ and $\eta_E \to 1$). Then it holds that

$$N_{ST} \le \frac{1}{1 - n_S^{-1}} \left[\frac{n_T}{n_{Bmax}} - 1 \right] + 3 . \tag{26}$$

If a provider has N_{STmax} summarization nodes available we can derive the minimum number of terminals in one group n_{Bmxm} from (26) and we get

$$n_{Bmxm} \ge \frac{n_T}{(N_{STmax} - 3)\left(1 - n_S^{-1}\right) + 1} . \tag{27}$$

When required parameter n_{Bmxm} obtained from (27) is smaller than optimum parameter n_{Bopt}, the point (n_{Bopt}, n_{Sopt}) will be used as the best value for the number of terminals in one group and for summarization nodes in one group respectively. Otherwise the new optimum value n_{Bopt} larger than n_{Bmxm} and n_{Sopt} will be searched. The results are separately compared in Fig. 6: and in Fig. 7:

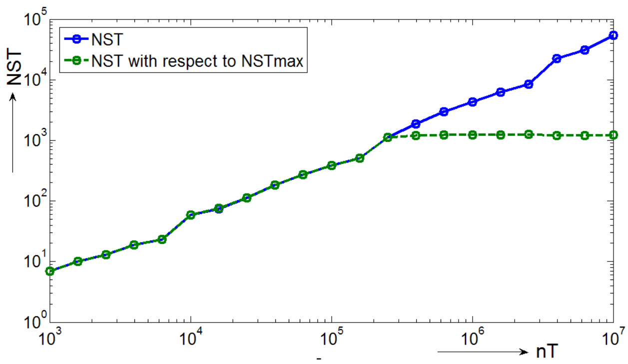


Fig. 6: The courses of required number of summarization nodes for different numbers of terminals (nT) without a and with respect to the demand on NSTmax = 1190.

to Contents, page 1

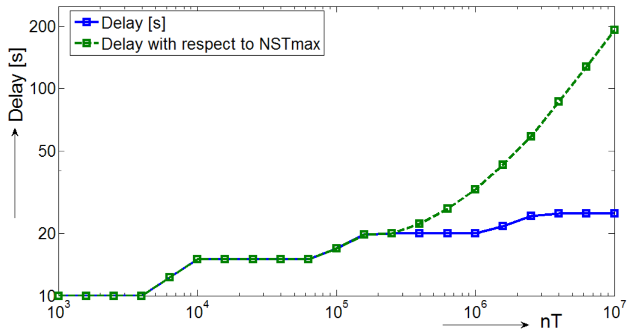


Fig. 7: The courses of minimum delay for different numbers of terminals (nT) without and with respect to the demand on NSTmax = 1190.

5. Conclusion

This article dealt with the problem of hierarchical data acquisition. The process of tree design was presented and some problems related to it were addressed like minimization of the total acquisition delay and the limited figure of summarization nodes. The delay optimum was found and tree parameters were derived. Influence of limited number of summarization nodes was considered and proved by simulations in Matlab environment. Separate paper will address the problem of end nodes (terminals) organisation according to their localities.

Acknowledgements

This paper a the research results were got with the support of Ministry of Education, Youth and Sports the project No. MSM0021630513.

References

[1] SCHULZRINNE, H.; CASNER, S.; FREDERICK, R.; JACOBSON, V. *RTP: A Transport Protocol for Real-Time Applications*. Internet Draft, IETF RFC3550, 2003. pp. 1-73

[2] HOLBROOK H.; CAIN, B. *Source-Specific Multicast for IP*. Internet Draft, IETF, 2004. pp. 1-17

[3] BHATTACHARYYA, S. *An Overview of Source-Specific Multicast (SSM)*. Request for Comments 3569, IETF, 2003. pp. 1-14

[4] CHESTERFIELD, J.; SCHOOLER, E., M. An extensible RTCP control framework for large multimedia distributions, In *Second IEEE International Symposium on Network Computing and Applications, 2003. NCA 2003*. 2003. pp. 351 – 358. ISBN 0-7695-1938-5.

[5] CHESTERFIELD, J.; OTT, J.; SCHOOLER, E., M. *RTCP Extensions for Single-Source Multicast Sessions with Unicast Feedback*. Internet draft, IETF, 2007. pp. 1-40

[6] EL-MARAKBY, R.; HUTCHISON, D. Scalability Improvement of the Real-Time Control Protocol (RTCP) Leading to Management Facilities in the Internet. In *Proceedings of Third IEEE Symposium on Computers & Communications*, ISCC. 1998. pp. 125 – 129. ISBN 0-8186-8538-7.

[7] CASTRO, M.; DRUSCHEL, P.; KERMARREC, A., M.; ROWSTRON, A., I., T. A large-scale and decentralized application-level multicast infrastructure. *IEEE Journal on Selected Areas in Communications*. 2002. Vol. 20, No. 8, pp. 1489-1499. ISSN 0733-8716.

[8] NOVOTNÝ, V.; KOMOSNÝ, D. Optimization of Large-Scale RTCP Feedback Reporting in Fixed and Mobile Networks. In *Proceedings of Third International Conference on Wireless and Mobile Communications, 2007. ICWMC '07*. March 2007, pp. 1–6. ISBN: 0-7695-2796-5.

[9] NOVOTNÝ, V.; KOMOSNÝ, D. Large-Scale RTCP Feedback Optimization. *Journal of Networks*, 2008, Vol. 3, No. 3, pp. 1-10. ISSN 1796-2056.

[10] KOMOSNÝ, D.; BURGET, R.; NOVOTNÝ, V. Tree Transmission Protocol for Feedback Distribution in IPTV Systems. In *Proceedings of the Seventh IASTED International Conference on Communication Systems and Networks*. Palma de Mallorca, Spain: International Association of Science and Technology for Development, 2008. pp. 1-7. ISBN 978-0-88986-758-1.

About Author

Vit NOVOTNY was born in 1969. He received his M.Sc. from Brno University of Technology in 1992, his Ph.D. degree received at the Brno University of Technology in 2001 in Electronics and communication technologies and he became the assistant professor at the same university in 2005, also in the area of “electronics and communication technologies”. In the past he did the research in the areas of non-filtering applications of switched capacitors and of the current and voltage conveyors. Current professional interests are mobile and packet data networks, their services and terminal equipment. Now he works with the Dept. of Telecommunications, Brno University of Technology.

to Contents, page 1

STUDY OF VIVALDI ALGORITHM IN ENERGY CONSTRAINT NETWORKS

Patrik MORAVEK[1], *Dan KOMOSNY*[1], *Pavel VAJSAR*[1], *Jaroslav SVEDA*[1], *Tomas HANDL*[1]

[1]Department of Telecommunication, Faculty of Electrical Engineering and Communication, Brno University of Technology, Purkynova 118, 612 00 Brno, Czech Republic

moravek@phd.feec.vutbr.cz, komosny@feec.vutbr.cz, xvajsa01@stud.feec.vutbr.cz

Abstract. *The presented paper discusses a viability of Vivaldi localization algorithm and synthetic coordinate system in general to be used for localization purposes in energy constraint networks. Synthetic coordinate systems achieve good results in IP based networks and thus, it could be a perspective way of node localization in other types of networks. However, transfer of Vivaldi algorithm into a different kind of network is a difficult task because the different basic characteristic of the network and network nodes. In this paper we focus on different aspects of IP based networks and networks of wireless sensors which suffer from strict energy limitation. During our work we proposed a modified version of two-dimensional Vivaldi localization algorithm with height system and developed a simulator tool for initial investigation of its function in ad-hoc energy constraint networks.*

Keywords

Energy, localization, simulations, synthetic coordinates, Vivaldi, wireless sensor networks.

1. Introduction

Wireless Sensor Networks (WSN) are an emerging network technology with promising perspective in the future. The networks consist of small low-cost, low-powered devices capable of sensing surrounding quantities and monitoring the environment. The devices called sensor nodes are equipped with radio interface for wireless communication. Standard that describes physical and link layer of such communication is IEEE 802.15.4 [1], for addressing, routing and other functions on higher levels Zigbee standard is often used [2].

WSN cover broad band of applications ranging from military projects (the historical origin of this technology) to medical surveillance including habitat monitoring, storehouse management or building automation.

In a lot of applications self-localization of network nodes is a high desirable feature. Sensed data without location information are meaningless. Moreover, other processes run in WSN can advantageously exploit the position knowledge for better and more efficient function. Geographical routing, hierarchical aggregation, multicast and data gathering can be mentioned as examples of such processes [3].

Because of a high variety of application in different fields and with different features, there are also a lot of different requirements on localization process. Some of the applications need a precise localization with an error less than 10 % but a coarse grained localization is sufficient for others. However, in general, WSN nodes are equipped with limited energy sources, and thus, each process in WSN should be energy aware.

One of the approaches increasing accuracy of localization in IP based network contains synthetic coordinate system Vivaldi algorithm. Our work discusses the viability of such approach in source limited networks such WSN.

The rest of the paper is organized as follows. Section 2 gives a brief overview of localization techniques and approaches, section 3 introduces synthetic coordinate systems and their representatives. The discussion of viability of synthetic coordinates and necessary modification of Vivaldi algorithm for use in WSN follow in section 4. New simulator designed for simulation of Vivaldi algorithm and its modification is described in section 5, which precedes the last summarizing section 6.

to Contents, page 1

2. Localization

The term “localization” relates to finding a position of an object in a defined area, generally. In IP based networks, the localization mainly means locating a station within a network. However, since a lot of applications provide surveillance, localization in WSN can also include a localization of an object, which is not a component of the network topology. This is called tracking and we do not consider it in this paper.

Localization protocols incorporate a localization algorithm to estimate the location of a sensor node without previous knowledge of its coordinates. The localization can be relative to the other nodes in a network or absolute in a determined coordinate system. If the network contains a certain percentage of nodes with a known position (called anchors), the unknown nodes (nodes having no knowledge of their position) use a certain measurement technique to estimate distances to these nodes and calculate their own position using determined localization algorithm. Anchor nodes can obtain their coordinates from GPS or by manual assignment. However, both approaches have their shortcomings either in higher energy cost or demanding initial process. The coordinate system of anchors is then applied to other nodes as well. The brief taxonomy of WSN localization is given in [4].

Localization algorithms require certain input information for the position determination. They work with information including mainly distances or angles. To obtain this information, specific measurement techniques are used. These techniques can be categorized into three main classes: RSS (received signal strength), TOA (time of arrival) and AOA (angle of arrival) based techniques [5]. Direct measurement is another technique, which is however impossible to use in the majority of applications.

RSS based method of distance estimation infers the distance from signal strength measured in a receiver. There are several signal propagation models that approximate the real radio channel and allow relating received signal strength to the distance between transmitter and receiver. It is an inexpensive and easy method of estimation in WSN since no extra hardware is required. However, several negative influences affect the measurement and cause estimation errors [6], [7].

Next category of measurements is based on measurement of signal propagation time. One-way measurement infers the distance between two nodes from the time of transmission and reception of packets. It requires precise a clock at each node and a complex time synchronization of all the nodes. To overcome this inconveniences, round-trip delay (RTT) can be calculated. The difference between the time of transmission and reception is measured at the same node, and thus, the synchronization is not necessary. However, a processing delay (to handle a packet) of the other node is included in the measured value. TDOA (time difference of arrival) is another method, which computes the position of a transmitter from the delay measured at several different nodes with a known position.

The last class of measurement techniques employs a system of angle measurements. If the unknown node knows at least the coordinates of two transmitting nodes and their directions, it is able to calculate its own position. For more detailed information about localization techniques and algorithms please refer to [5].

3. Synthetic Coordinate Systems

To improve localization accuracy in IP based networks, the problem of the distance estimation between stations was transferred into an artificial multidimensional coordinate system. Generally, RTT measurement was performed to estimate the distance between two stations. There is no need to measure the RTT between each pair of stations in localization algorithm based on a synthetic coordination system. Instead, the RTT value between two stations is estimated from their known coordinates in a predefined synthetic coordinate system. The RTT value refers to a distance between the stations in the coordinate system. Provided that data networks work ideally (there are no delays), a geographical coordinate system with longitude and latitude would be the appropriate choice as a coordinate system for localization purposes. Unfortunately, this condition is not accomplished in any current data network – packets are transmitted via more direct links, delayed at intermediate nodes, etc. Therefore, it is not possible to use a simple 2D coordinate system for RTT value prediction of network nodes. As a result, new artificial coordinate systems were proposed to meet conditions in IP based networks. There is no limitation in the number of dimensions or the type of coordinate system. Besides the standard Euclidian system, other coordinate systems (spherical, toroid, hyperbolic) were investigated (for more details see [8]). However, the Euclidian coordinate system is used the most because it offers the most suitable possibilities for this purpose.

There are proposed several algorithms using synthetic coordinate system (GNP, Lighthouse system, Vivaldi etc.). GNP [9] is a centralized algorithm with reference stations, which form a matrix of distances between themselves in the first phase and the rest of nodes is localized in the following step.

Lighthouse system presented in [10] uses reference station as well (called lighthouses). Contrary to GNP, it uses simpler mathematical operations and features more scalability thanks to employing recently localized nodes into a reference nodes infrastructure in each iteration step.

to Contents, page 1

3.1. Vivaldi Algorithm

The Vivaldi algorithm proposed by Dabek et al. in [11] is a favorite localization algorithm used to obtain the position of stations in a network using the synthetic coordinate system. The algorithm uses synthetic coordinates with Euclidian distances; two standard dimensions and one extra dimension called height are defined in the new coordinate system. Communication delay in an access network is covered by the third dimension (as the main purpose of this dimension) while delay in a distribution network is expressed by coordinates.

The new coordinate system, 2D Euclidian system with height, is described by the following equations:

$$x - y = (x_1 - y_1, x_2 - y_2, x_h + y_h)\,, \tag{1}$$

$$\|x\| = \sqrt{{x_1}^2 + {x_2}^2} + x_h\,, \tag{2}$$

$$ax = (ax_1, ax_2, ax_h)\;. \tag{3}$$

Coordinates of nodes are taken as vectors in the whole following text. The difference between standard 3D, 2D system and 2D coordinate system with height is depicted in Fig. 1.

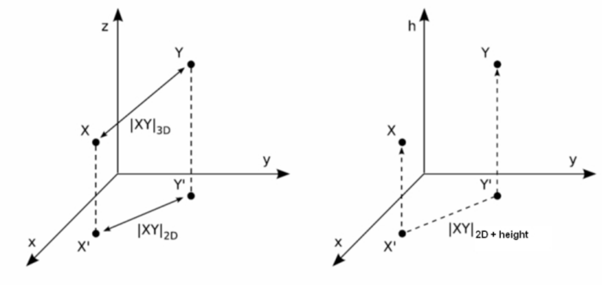


Fig. 1: Distance in 2D, 3D and 2D with height coordinate system.

The Vivaldi algorithm is a distributed and decentralized algorithm working without any infrastructure (such as reference stations). All nodes are equivalent in the system.

Finding node coordinates that minimize the error in predicted round-trip latency between arbitrary two nodes in a network is the basic principle of the Vivaldi algorithm. The idea of the algorithm comes from the analogy to a physical mass-spring system described in [12].

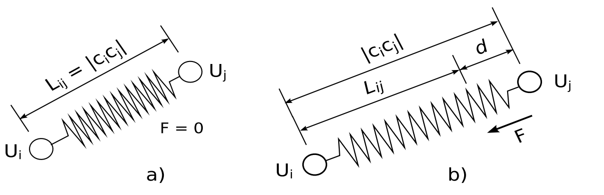


Fig. 2: Spring analogy to predicted latency; a) errorless prediction b) prediction with error.

As a spring tends to maintain its length with minimum energy, distances between nodes in a network are set such that minimum predicted latency error is achieved. Findings the energy minimum in the system of springs corresponds to finding the minimum error in the position estimation. Searching for node position is simulated as a movement of nodes connected with springs, see Fig. 2.

Nodes in a network are moved in a way in order to minimize the error function E:

$$E = \sum_i \sum_j \left(L_{ij} - \left\| \vec{\boldsymbol{c}}_i - \vec{\boldsymbol{c}}_j \right\| \right), \tag{4}$$

where L_{ij} is the real latency between nodes i and j and $\vec{\boldsymbol{c}}_i$ and $\vec{\boldsymbol{c}}_j$ are their coordinates in the synthetic coordinate space. This equation corresponds to a spring between i and j nodes with a length of L_{ij}.

The principle of minimizing the error function is derived from the impact of a spring placed between two nodes. According to the Hook law

$$\vec{\boldsymbol{F}} = -k\vec{\boldsymbol{d}}\,, \tag{5}$$

the stretched or compressed spring affects the surrounding nodes, which are linked by the spring, by the force $\vec{\boldsymbol{F}}$ in the opposite direction to this stretch or compression. This force is proportional to the length of the stretch (compression) described by the vector $\vec{\boldsymbol{d}}$ and the spring constant k. In the Vivaldi algorithm, the force $\vec{\boldsymbol{F}}_{ij}$ affecting the nodes i and j is

$$\vec{\boldsymbol{F}}_{ij} = \left(L_{ij} - \left\| \vec{\boldsymbol{c}}_i - \vec{\boldsymbol{c}}_j \right\| \right) \vec{\boldsymbol{u}}\left(\vec{\boldsymbol{c}}_i - \vec{\boldsymbol{c}}_j \right), \tag{6}$$

Where $\vec{\boldsymbol{u}}\left(\vec{\boldsymbol{c}}_i - \vec{\boldsymbol{c}}_j \right)$ is the unit vector with the same direction as the vector $\vec{\boldsymbol{c}}_i - \vec{\boldsymbol{c}}_j$.

The resultant vector $\vec{\boldsymbol{F}}_i$ of a node i is the summation of partial vectors $\vec{\boldsymbol{F}}_{ij}$ affecting the node i (from all springs connected to the node i). In the one iteration step of the computational process the time interval δ of the affecting force $\vec{\boldsymbol{F}}_i$ is considered. The node i is subsequently moved to a new position $\vec{\boldsymbol{c}}_i$ in the direction of force $\vec{\boldsymbol{F}}_i$ according to the equation

$$\vec{\boldsymbol{c}}_i = \vec{\boldsymbol{c}}_i + \delta\, \vec{\boldsymbol{F}}_{ij}\,, \tag{7}$$

Since the Vivaldi algorithm is a decentralized localization algorithm, the presented idea is performed at all nodes in a network. Each node then individually simulates its movement. The direction and the length of movement are computed from the latency values L_{ij} and

to Contents, page 1

coordinates $\vec{\boldsymbol{c}}_j$ received from nodes *j*.

The decentralized character of the algorithm means also that the coordinates received do not have to be reliable. The node with the coordinate $\vec{\boldsymbol{c}}_j$ can be, for example, a new node at the beginning of the localization process or a node that cannot determine its position for whatever reason, and its coordinate oscillates.

The negative impact of the situations described is reduced in the Vivaldi algorithm with an adaptive timestep by assigning a specific error e_j to each node. This error is sent by node with its coordinate.

The complete process of algorithm for node *I* is described in five steps below.

1. the weight *w* is computed from the estimated errors in coordinate calculation at the local node *i* and a distant node *j*

$$w = \frac{e_i}{e_i + e_j}. \tag{8}$$

2. the relative error e_s of latency measurement calculation

$$e_s = \frac{L_{ij} - \left\| \vec{\boldsymbol{c}}_i - \vec{\boldsymbol{c}}_j \right\|}{L_{ij}}. \tag{9}$$

3. the weighted moving average of local error e_i is updated

$$e_i = e_s c_e w + e_i (1 - c_e w). \tag{10}$$

4. timestep δ calculation

$$\delta = c_c w. \tag{11}$$

5. the node coordinate is updated

$$\vec{\boldsymbol{c}}_i = \vec{\boldsymbol{c}}_i + \delta (L_{ij} - \left\| \vec{\boldsymbol{c}}_i - \vec{\boldsymbol{c}}_j \right\|) \vec{\boldsymbol{u}} (\vec{\boldsymbol{c}}_i - \vec{\boldsymbol{c}}_j). \tag{12}$$

One modification of the Vivaldi algorithm implements also the timestep δ adapted by multiplication by the constant c_c and the weight *w* (step 4). The weight *w* depends on both local error e_i and distant error e_j. If the error e_j is greater related to local error e_i the weight *w* is smaller and the node *j* has small relevance in position calculation. On the contrary, if the error e_j is small, the weight *w* is close to one and the node *j* impacts on position calculation significantly.

The error e_i in coordinate estimation is calculated in step 3 as a weighted moving average of relative latency error e_s. The value of this average can be changed by a tune constant c_e. If the constant c_e is close to one, the error e_i is affected the most by the current error e_s. With decreasing value of c_e the previous value of e_i plays a more significant role in the calculation.

Besides the above described algorithm, two simpler variants exist [11]. The main difference is that the timestep δ is a constant or a slowly decreasing value instead of dynamically adapting in these modifications.

4. Synthetic Coordinates in WSN

Since synthetic coordinate systems were proposed for IP based networks, there are several inconveniences rising from their usage in WSN. Algorithms described in the previous section are very demanding; especially from the energy point of view. We have to be always aware of strict energy constraints relating with WSN technology. Sensor nodes are relatively simple devices with weak microcontroller and very limited energy source. And provided that we use synthetic coordinate system, we affect both. There is higher computation cost and because of the frequent communication, the energy consumed by a radio part also increases. Moreover, system management and control require certain amount of energy too. On the other hand, the synthetic coordinate system offers indisputable advantages. Decentralized feature of a localization algorithm means that the system is less vulnerable to system collapse because of node dysfunction or local error. There is an option to start localization without anchor nodes and form a completely relative map. But mainly, it provides more accurate position estimation based on the cooperation of all nodes. The synthetic coordinate algorithms are able to eliminate or minimize error caused by measurement methods, which is a serious problem in the range based localization. The optimization of accuracy is based on iterative approximation. However, this means an undesirable increase of energy cost, since each iteration requires updated information about the position of other nodes and a new measurement of the distance.

All the mentioned facts infer that we have to accept a certain trade-off considering the use of synthetic coordinate systems in WSN. Also, application requirements set the important conditions and limitations. Therefore, we proposed following modification of Vivaldi algorithm to adapt it for wireless sensor networks. The modification is called EAVA (Energy Aware Vivaldi Algorithm) and we will refer to it in the following text.

First, we decided to use two dimensional system with height (2D+h) with possible extension to 4D+h system, which can have better results (as stated in [13]). In IP based networks *h* is a positive value since it represents delay between two stations. In WSN, *h* can be related to general error caused by measurement method. Thus, it can be either positive or negative.

The distances are derived from RTT measurements in IP based networks. However, this is

to Contents, page 1

hardly possible in WSN. Time measurement requires precise time synchronization, which means a precise clock embedded in each device. Moreover, time synchronization process and its control is a difficult task and additional energy costs. In low-cost applications, mostly RSS based distance estimation is used. So, we recommend using the RSS measurement for the distance estimation instead of time based measurements. In proposed simulations with EAVA we consider RSS measurement as well.

The initial setting of a network depends on the presence of anchor nodes. If there are some, they can be either equipped with GPS receiver to set the coordinates or set manually. Then, the third coordinate *h* states for the error of GPS estimation. The other two coordinates relate to the standard 2D geographical system. With this initial setting of anchor nodes, triangulation or maximum likelihood method is performed to obtain a rough position estimation of all the nodes.

Provided that there are no anchor nodes in the network, all the nodes are set with certain determined initial coordinates (such as x_i=0, y_i=0, h_i=0, for example). The localization process then starts from the very beginning and there is naturally slower convergence requiring more iteration steps, which consumes subsequently more energy.

To save energy during the communication before each iteration step, only the RSS measurement and the communication with neighbours is performed. In case of high node degree (number of neighbours), only a subset of neighbours can be involved.

The energy consumption is the most crucial parameter of the localization, highly dependent on the number of iterations and the speed of convergence. Therefore, the setting of a shift constantΔ analogical to the timestep δ is highly important. However, there are other conditions and parameters, which considerably influence the convergence, and thus, the energy depletion and they have to be investigated. For parameters adjustment purposes certain simulations were proposed.

5. Simulations

The adjustment of protocols developed for a certain kind of networks and their transfer to an environment with totally different main features is always a challenging process. It is difficult to predict functionality and reliability of a protocol under different conditions. Although, a certain protocol works well in IP networks, it can totally fail in WSN. The energy depletion is the most problematic in this case, which is not considered in networks of mains-operated stations. Therefore, it is necessary to run certain simulations to verify the viability of a transferred protocol in new environment. Simulations can answer the question if it is reasonable to use the protocol without changes, with changes or if the protocol is totally inapplicable and a new one should be proposed.

For the purposes of the simulation of synthetic coordinate algorithms Vivaldi and EAVA, we proposed and developed a new simulation tool [14], [15]. The simulator is a JAVA based application implementing Vivaldi algorithm and its modified versions. The base coordinate system is 2D with height but can be easily upgraded to 4D with height system. The main simulator window can be seen in Fig. 3.

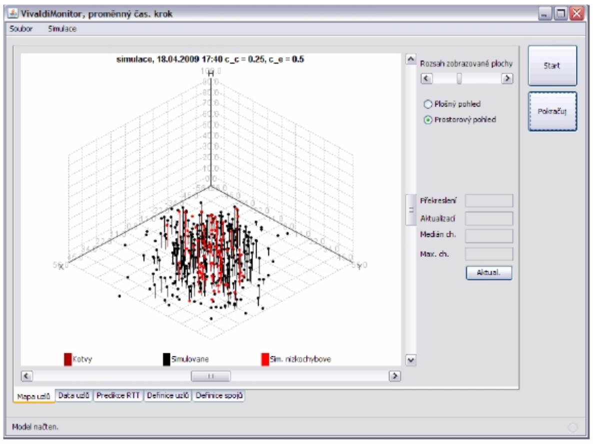


Fig. 3: Graphic user interface of developed simulation tool.

The simulator displays the convergence process of localization by updating the main window. There are three types of network nodes distinguished in the figure; anchors, unknown nodes with position error under threshold (0,15 by default) and nodes with higher position error.

The other modules of the simulator offer the graphical representation of position and error evolution. The example of a position convergence simulation is depicted in Fig. 4.

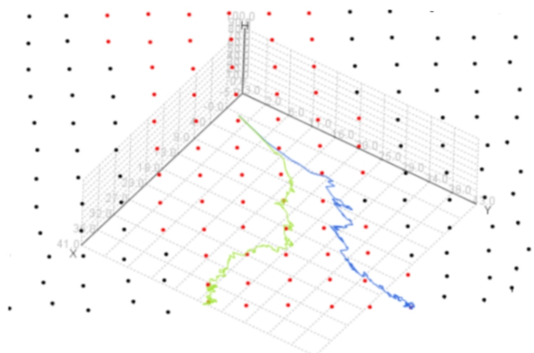


Fig. 4: Convergence of position of two nodes.

Another example of a simulator output can be seen in Fig. 5, Fig. 6 and Fig. 7. They show the two dimensional area of deployed nodes after simulation.

to Contents, page 1

First, the real node deployment is defined (or imported) in the simulator. Then, the distance metrics are given to the nodes (defined or imported). When the simulator runs the Vivaldi or EAVA algorithm it successively updates the picture of node deployment. The figures present the result picture after 50, 700 and 1500 iteration steps respectively. The black marks in the figures represent the nodes, which are still far away from their real location, it means, they have a large position error. On the other hand, the red mark represents a node, whose position error is already under the defined threshold. The more iteration the simulator performs, the higher probability of successful localization (low position error) is.

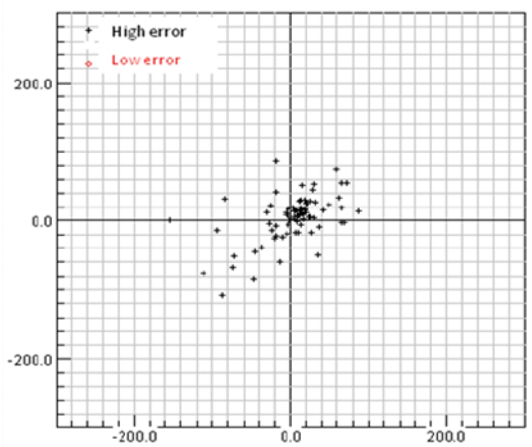


Fig. 5: Node deployment after 50 iterations of Vivaldi algorithm with adaptive iteration step.

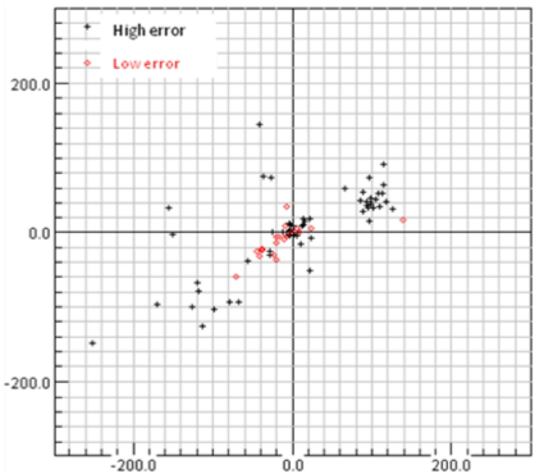


Fig. 6: Node deployment after 700 iterations of Vivaldi algorithm with adaptive iteration step.

The Fig. 8 depicts the sample simulation result where the error is evaluated. We have both absolute error and relative error at disposal. Since the original Vivaldi algorithm uses a time distances between nodes for position estimation and for localization optimization the absolute error is expressed in ms. The absolute error is a difference between the RTT of nodes at their real locations and the RTT of the nodes in a new network arrangement calculated by the localization algorithm.

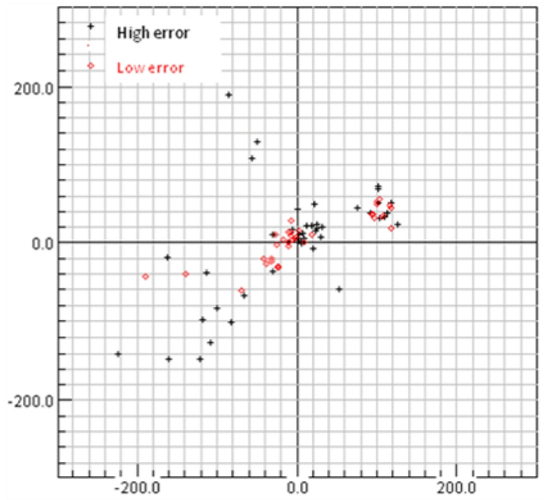


Fig. 7: Node deployment after 1500 iterations of Vivaldi algorithm with adaptive iteration step.

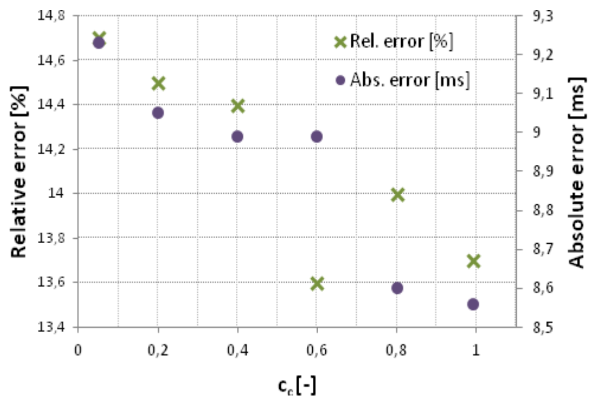


Fig. 8: Relative and absolute error of sample scenario after the simulation in dependency on the parameter c_c of EAVA.

Besides that the simulator incorporates the embedded editor of simulated topologies with import and export functions. Moreover, the various quantities and their evolution are presented in clearly arranged graphs with the possibility of subsequent export of all the data into a Matlab environment. History of coordinates is also a useful feature.

6. Conclusion and Future Work

Localization in WSN is a challenging task and there is a broad variety of approaches to this topic. Because of the particularity of these networks it is very difficult to simply implement protocols from different networks. However, although it is impossible to transfer protocols directly, it is promising to use at least some their features, which allowed their successful deployment in IP based

to Contents, page 1

networks. Therefore, we proposed modified version of Vivaldi algorithm called EAVA, which is adjusted to WSN. The protocol considers RSS measurement for distance estimation and strictly controls the energy consumption during localization. The main idea is to exploit cooperation of nodes based on mass spring principle and at the same time use as little energy as possible. In IP based networks the Vivaldi algorithm features promising results but at the cost of high communication. This is not acceptable in WSN, so radical change has to be done.

For the simulation purposes, we proposed and developed a new simulator tool VivaldiMonitor. It implements successfully Vivaldi algorithm and its modification into static networks. It offers modularity and elaborated graphical output with user-friendly interface with data export and import option. The next phase is devoted to the development of a library implementing EAVA and the modification of the simulator for WSN specific features with development of energy module to control energy depletion during localization.

Acknowledgements

This work was supported by Ministry of Education of the Czech Republic project No. MSM0021630513, Ministry of Industry and Trade of the Czech Republic - project No. MPO FR-TI2/571.

References

[1] IEEE Standard for Information Technology Part 15.4: *Wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs)*, IEEE Std 802.15.4-2006 (Revision of IEEE Std 802.15.4- 2003), pp. 1–305, 2006.

[2] *Zigbee Alliance*, [online]. Available at WWW: <www.zigbee.org.>.

[3] AKYILDIZ, F., I.; SU, W.; SANKARASUBRAMANIAM, Y.; CAYIRCI, E. Wireless sensor networks: A survey. *Computer Networks*, 2002, Vol. 38, No. 4, pp. 393–422. ISSN 1389-1286.

[4] HUANG, W.; WANG, Y.; GUAN, H. The current situation and prospect of localization in wireless sensor network. In *Computer Science and Engineering, 2009. WCSE '09. Second International Workshop*, 28-30 Oct. 2009, pp. 483-487. ISBN 978-0-7695-3881-5.

[5] MAO, G.; FIDAN, B.; ANDERSON, B., D., O. Wireless sensor network localization techniques. *Computer Networks*, 2007, Vol. 51, No. 10, pp. 2529-2553. ISSN 1389-1286.

[6] MORÁVEK, P.; KOMOSNÝ, D.; SALA, D., G.; GUILLÉN, A., L. Received signal strength uncertainty in energy-aware localization in wireless sensor networks. In *Environment and Electrical Engineering (EEEIC), 2010 9th International Conference on*, 2010. pp. 538-541, ISBN 978-1-4244-5370-2.

[7] MORAVEK, Patrik; KOMOSNY, Dan; SIMEK, Milan; JELINEK Mojmir; GIRBAU, Sala, David LAZARO, Guillen, Antoni. Signal Propagation and Distance Estimation in Wireless Sensor Networks. In *The 33rd International Conference on Telecommunication and Signal Processing, TSP 2010*. 2010. pp. 35-40, ISBN 978-963-88981-0-4.

[8] SHAVITT, Y.; TANKEL, T. On the curvature of the Internet and its usage for overlay construction and distance estimation. In *INFOCOM 2004 - Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*. IEEE, 2004. Vol. 1, ISBN 0-7803-8355-9.

[9] NG, E.; HUI, Z. Predicting Internet network distance with coordinates-based approaches. In *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies*. Proceedings. IEEE, 2002. Vol. 1, pp. 170 – 179, ISBN 0-7803-7476-2.

[10] PIAS, M.; CROWCROFT, J.; WILBUR, S.; HARRIS, T.; BHATTI, S. *Lighthouses for Scalable Distributed Location*. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Berlin, Heidelberg, 2003. pp. 278-291, ISBN 978-3-540-40724-9.

[11] DABEK, F.; COX, R.; KAASHOEK, F.; MORRIS, R. Vivaldi: A decentralized network coordinate system. In *ACM SIGCOMM Computer Communication Review*. Association for Computing Machinery, 2004. Vol. 34, No. 4, pp. 15-26. ISSN 0146-4833.

[12] FRUCHTERMAN, T.; REINGOLD, E. *Graph Drawing by Force-directed Placement*. Software - Practice and Experience (SPE), November 1991, ISSN 1097-024X.

[13] LEDLIE, J.; GARDNER, P.; SELTZER, M. Network Coordinates in the Wild In *Proceedings of 4th USENIX Symposium on Networked Systems Design & Implementation 2007*. [online] Cambridge (USA): Harvard School of Engineering and Applied Sciences and Aelitis, 2007. pp. 299–311, Available at WWW: <http://www.eecs.harvard.edu/ syrah/nc/wild07.pdf>.

[14] MORAVEK, Patrik; KOMOSNY Dan; SIMEK Milan; SVEDA, Jaromir; HANDL, Tomas. Vivaldi and other localization methods. In *Proceeding of 32nd International Conference on Telecommunications and Signal Processing, TSP 2009*. 2009. pp. 1-5. ISBN 978-963-06-7716-5.

[15] MORAVEK, Patrik; KOMOSNY Dan; BURGET Radim; SVEDA, Jaromir; HANDL, Tomas. Study and Performance of Localization Methods in IP Based Networks: Vivaldi Algorithm. *Journal of Network and Computer Applications*, 2010. Vol. 34, No. 1, pp. 1-17. ISSN 1084-8045.

About Authors

Patrik MORAVEK is a doctoral student in the Department of Telecommunication at the Brno University of Technology. He gained his Master’s degree at the same university in 2008. His study specialization is TCP/IP stack and communication in ad-hoc networks. Nowadays, he is focused on wireless sensor network technologies and especially on the energy efficiency of radio communication and localization process.

Dan KOMOSNY received his Ph.D. degree in Teleinformatics in 2003. His research is focused on wireless sensor networks, IP networks and network

to Contents, page 1

operating systems. He gives lectures in courses dealing with Linux operating systems and communication in IP networks. He is an internal member of technical program committee of IEEE and IASTED conferences.

Pavel VAJSAR gained his Master degree at the Faculty of Electrical Engineering and Communication Technologies at the Brno University of Technology. Nowadays, he is a doctoral student at the same faculty. His major interest covers wireless sensor networks and their monitoring. Moreover, his research contributes to routing in MANET networks considering quality of service. He is also an OPNET modeler expert.

Jaroslav SVEDA gained his Master degree at the Faculty of Electrical Engineering and Communication Technologies at the Brno University of Technology in 2009. His interest covers network technologies and simulations.

Tomas HANDL gained his Master degree at the Faculty of Electrical Engineering and Communication Technologies at the Brno University of Technology in 2009. His field of expertise is mostly in IP based networks.

to Contents, page 1

INNOVATION OF METHODS FOR MEASUREMENT AND MODELLING OF TWISTED PAIR PARAMETERS

Lukas CEPA[1], *Milos KOZAK*[1], *Jiri VODRAZKA*[1]

[1]Department of Telecommunication Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Technicka 2, 166 27 Prague, Czech Republic

cepaluka@fel.cvut.cz, milos.kozak@fel.cvut.cz, vodrazka@fel.cvut.cz

Abstract. *The goal of this paper is to optimize a measurement methodology for the most accurate broadband modelling of characteristic impedance and other parameters for twisted pairs. Measured values and theirs comparison is presented in this article. Automated measurement facility was implemented at the Department of telecommunication of Faculty of electrical engineering of Czech technical university in Prague. Measurement facility contains RF switches allowing measurements up to 300 MHz or 1 GHz. Measured twisted pair's parameters can be obtained by measurement but for purposes of fundamental characteristics modelling is useful to define functions that model the properties of the twisted pair. Its primary and secondary parameters depend mostly on the frequency. For twisted pair deployment, we are interested in a frequency band range from 0,1 to 100 MHz.*

Keywords

Twisted pair, British telecom model, measurement method.

1. Introduction

The secondary parameters of a twisted pair may be modeled directly, as implements the Deutsche Telekom, or it can model primary parameters of a twisted pair. This way of modelling has chosen British Telecom (BT). Primary parameters, which are frequency dependent, can be designed for a limited frequency range and limited accuracy. There are several empirical models based on the frequency band [2]. 9-parameters BT model [1] can be used as model of primary parameters up to 100 MHz. 9-parameters BT model is given by following equations:

$$R(f)=\sqrt[4]{r_0^2+a\cdot f^2}\,, \tag{1}$$

where r_0 is the DC resistance, a is a parameter which characterizes the slope of the frequency response.

$$L(f)=\frac{l_0+l_\infty\cdot\left(\dfrac{f}{f_m}\right)^b}{1+\left(\dfrac{f}{f_m}\right)^b}\,, \tag{2}$$

where l_0 is inductance at zero frequency limiting, l_∞ is inductance at the highest frequencies of the band, fm and b are parameters that characterize the threshold between low and high frequencies.

$$C(f)=c_\infty+c_0\cdot f^{-c_g}\,, \tag{3}$$

where c_0 is capacity at the highest frequencies in the band, c∞ is capacity at zero frequency limiting and cg is parameter characterizing the shape of frequency-dependent waveform.

$$G(f)=g_0\cdot f^{g_e}\,, \tag{4}$$

where g_0 is conduction at low frequency and ge is coefficient increasing towards to high frequency.

A certain problem at high frequencies the impedance inhomogeneity of transmission line that cause partial reflections and then ripple on attenuation frequency dependence of twisted pair [5].

2. Measurement Methods

The twisted pair was chosen in order of measurement methods comparison. Method with internal and external bridge was chosen.

2.1. Internal Bridge Method

The internal bridge method uses North Hill 032BF balun

to Contents, page 1

(transformer for conversion from balance to unbalance port and vice versa) which acts as an internal measuring bridge between unbalance port with termination 50 Ω and balance port with termination 100 Ω. The balun works with frequencies from 100 kHz to 100 MHz. Block diagram is depicted in Fig. 1.

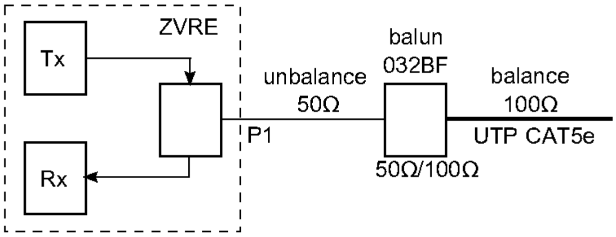


Fig. 1: Block diagram for measurement method with internal bridge North Hill 032BF.

The internal method is based on obtaining of return loss (RL, S11 parameter) at the input of Vector Network Analyzer ZVRE. Measured twisted pair is terminated by short, open and matched impedance. The input impedance of the twisted pair is given by equations:

$$Z_c(f) = Z_0 \cdot \frac{1+\Gamma}{1-\Gamma}, \tag{5}$$

where Z_{ref} is reference 100 Ω impedance of the twisted pair and Γ is reflection coefficient of the twisted pair terminated by short, open or matched impedance.

The characteristic impedance of the twisted pair Zo is given by equations:

$$Z_0(f) = \sqrt{Z_{x,open}(f) \cdot Z_{x,short}(f)}, \tag{6}$$

where $Z_{x,open}$ and $Z_{x,short}$ are measured impedances obtained from (5), where open and short termination was considered.

2.2. External Bridge Method

The external bridge method uses North Hill 0319NA which acts as an external measuring bridge between unbalance ports with termination 50 Ω and balance port with termination 100 Ω for twisted pair. Measuring bridge works with frequencies from 100 kHz to 100 MHz. Block diagram is depicted in Fig. 2.

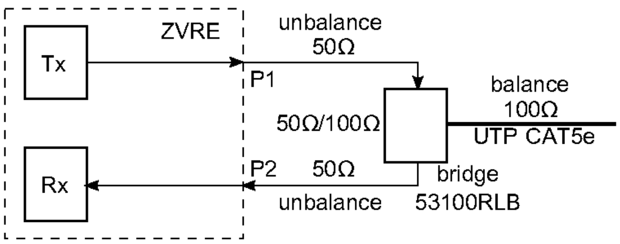


Fig. 2: Block diagram with external measuring bridge North Hill 53100RLB.

The external bridge method is based on obtaining of the forward transmission coefficient (S21 parameter). The external measuring bridge is used in the frequency range from 1 MHz to 100 MHz. Measured twisted pair is terminated by short, open and matched impedance. In order of proper results all measured values from Vector Network Analyzer ZVRE have to be corrected. Correction coefficient is obtained after removing of the twisted pair and leaving the port open. Revised value of return loss is given by:

$$\Gamma_{cor} = \frac{\Gamma}{\Gamma_{open}}, \tag{7}$$

where Γ is the reflection coefficient of the twisted pair terminated by short, open or matched impedance and Γopen is reflection coefficient of an external measuring bridge terminated by open.

Regarding of manual correction the return loss is obtained (S11 parameter) instead of forward transmission coefficient (S21 parameter). It is also necessary to determine the standardized characteristic impedance of the twisted pair, because of the transformation from 50 Ω to 100 Ω twisted pair according to the equation:

$$Z'_{norm}(f) = Z_0 \cdot \frac{1+\Gamma_{ref}}{1-\Gamma_{ref}}, \tag{8}$$

where Γ_{ref} is reflection coefficient of the external measuring bridge terminated by the reference resistor (100 Ω).

The input impedance of the twisted pair Z_x is given by the Eq. (9). The characteristic impedance of the twisted pair Z_0 is given by the Eq. (10).

$$Z(f) = Z'_{norm} \cdot \frac{1+\Gamma_{cor}}{1-\Gamma_{cor}}, \tag{9}$$

$$Z_0(f) = \sqrt{Z_{x,open}(f) \cdot Z_{x,short}(f)}, \tag{10}$$

where $Z_{x,open}$ and $Z_{x,short}$ are measured impedances obtained from (9), where open and short termination is considered.

3. Measurement Workplace

At the Department of Telecommunications Technology of Faculty of Electrical Engineering of CTU in Prague was implemented measurement facility which allows automatic measurement. This facility is depicted in Fig. 3.

The facility consists of:

- Vector Network Analyzer ZVRE firm ROHDE SCHWARZ,

to Contents, page 1

- Agilent 3499B 2-Slot Switch/Control Mainframe,
- 2x Agilent 44478A Dual 1x4 RF Multiplexer Module (1,3 GHz, 50 Ω). Dual 1-to-4 multiplexers provide bidirectional switching of signals from DC to 1,3 GHz,
- Baluns 0319NA firm North Hill, with the frequency,
- bandwidth from 100 kHz to 300 MHz,
- 50 Ω cables,
- Measured UTP cable.

The goal of a measuring facility is a faster and more accurate measurement of the secondary parameters, input impedance and crosstalk at the far end FEXT (Far End Cross-talk). Switching of RF (Radio Frequency) switches and processing of values is controlled by computer. Switching matrix also performs termination of unused pairs.

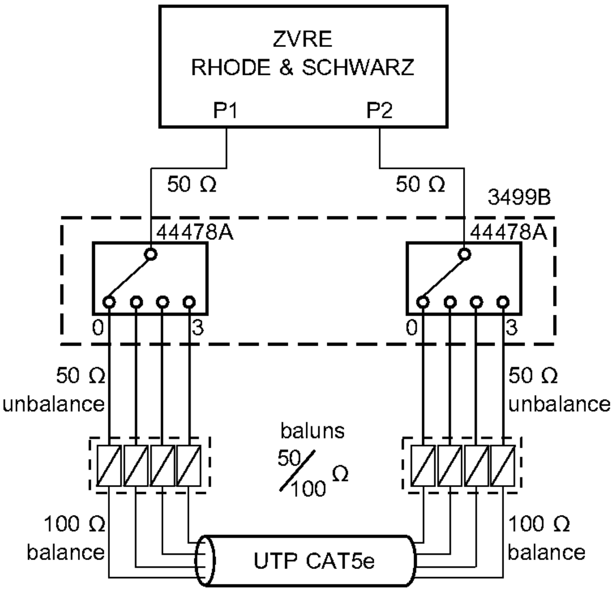


Fig. 3: Block diagram of measuring facility of the Department of Telecommunications Technology.

The detailed diagram of measuring facility is on Fig. 4 (for attenuation, FEXT and RL measurement) and Fig. 5 (for NEXT measurement).

Measuring facility contains additionally network analyzer ZVRE Rohde & Schwarz connected to port CH00 and CH10 of RF switches. Ports are equipped by internal termination resistors which are important for FEXT measurements. Red dashed lines depicted on Fig. 4 are states of switches for NEXT (Near End Cross-talk) measurements which are executed between port CH00 and CH10. All four channels of measuring facility are connected to near and far end of tested cable using adaptor for termination and balancing (coaxial interface of 50 Ω to 100 Ω for symmetrical cable).

Measuring facility setup for NEXT measurement is depicted on Fig. 5. This setup is special because allows to change channels for generator and receiver at the same end of tested cable. This operation is carried out thanks to internal ports of the switch (CH00-03), which is primary designed for termination of the cable. The connection is made by four cables to second switch (at CH10-13). The crosstalk modelling is presented in detail in [4].

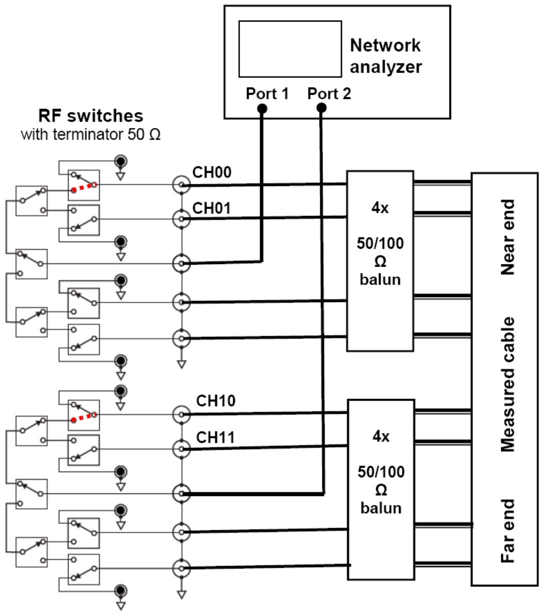


Fig. 4: Detailed block diagram of measuring facility setup for attenuation, FEXT and RL measurements.

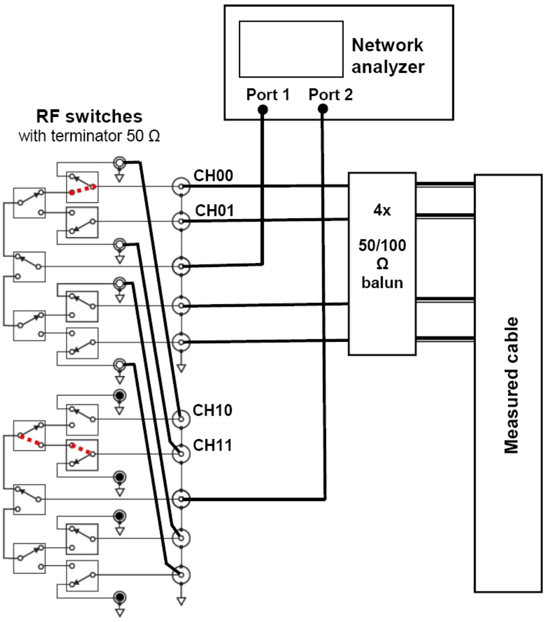


Fig. 5: Detailed block diagram of measuring facility setup for NEXT measurement.

to Contents, page 1

4. Measurement Results

The measured secondary parameters of each pair (transformed at 1 km cable length) are depicted in Fig. 6.

Figure 7 depicts a difference between waveforms of $|Z_0|$ measured using the methods above. Method with an external bridge has more stable waveform of the values $|Z_0|$ that is probably caused by attenuation reflection coefficient between ZVRE and external measuring bridge.

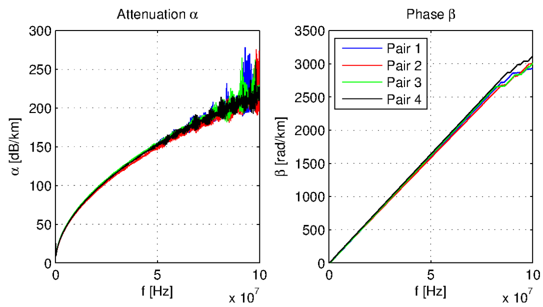


Fig. 6: Waveforms of the secondary parameters of each pair CAT5 UTP cable.

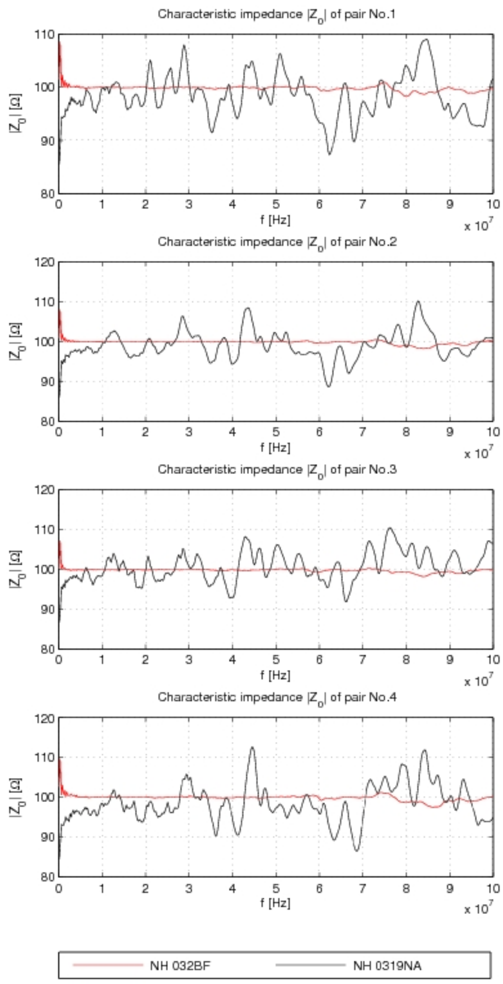


Fig. 7: Waveforms of |Z0| for pair in UTP cable.

The primary parameters (Fig. 8) were obtained using Eq. (11), (12), (13), (14) from secondary parameters.

$$R = \alpha \cdot \Re(Z_0) - \beta \cdot \Im(Z_0), \tag{11}$$

$$G = \frac{\alpha \cdot \Re(Z_0) + \beta \cdot \Im(Z_0)}{\Re(Z_0^2) + \Im(Z_0^2)}, \tag{12}$$

$$\omega \cdot L = \beta \cdot \Re(Z_0) + \alpha \cdot \Im(Z_0), \tag{13}$$

$$\omega \cdot C = \frac{\beta \cdot \Re(Z_0) - \alpha \cdot \Im(Z_0)}{\Re(Z_0^2) + \Im(Z_0^2)}. \tag{14}$$

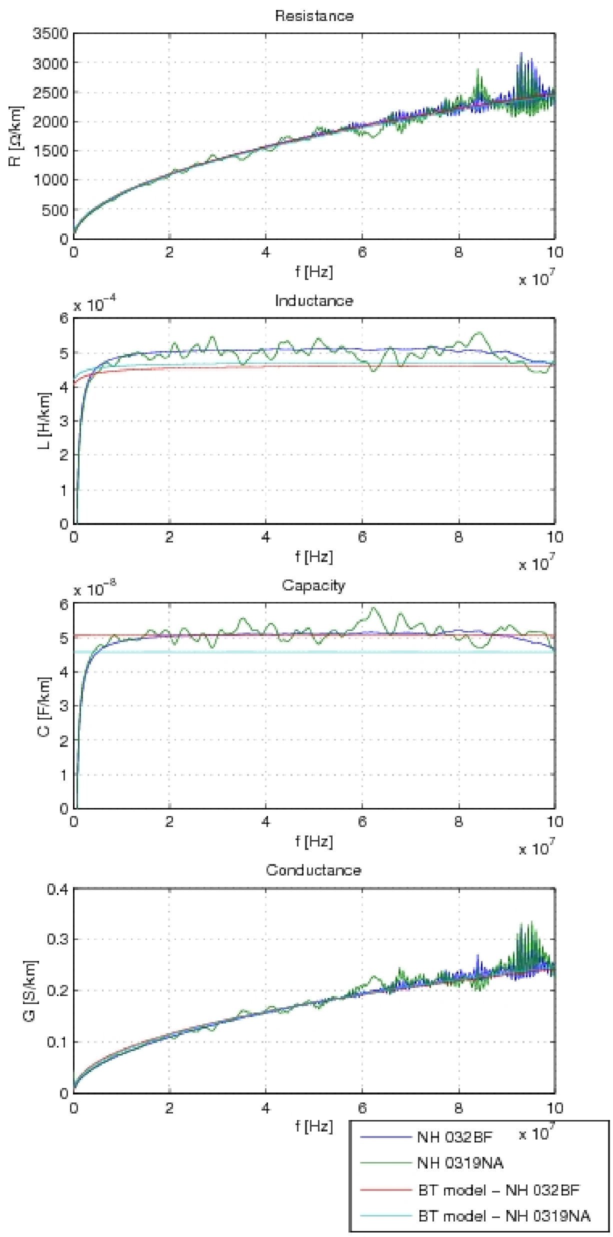


Fig. 8: Waveforms of the primary parameters of pair No. 1.

to Contents, page 1

5. Modelling Parameters

The BT model parameters were identified by Matlab from waveforms of Fig. 8. Tab. 1 contains a summary of identified BT model parameters for the two measurement methods; see Eq. (1) to (4).

Tab.1: Summarized parameters of BT model of two methods.

Parameters	NH 032BF	NH 53100RLB
a	$3{,}758\cdot10^{-3}$	$3{,}542\cdot10^{-3}$
r_o	90,655 Ω/km	90,655 Ω/km
l_o	$404{,}434\cdot10^{-6}$ H/km	$415{,}632\cdot10^{-6}$ H/km
l_∞	$464{,}784\cdot10^{-6}$ H/km	$473{,}367\cdot10^{-6}$ H/km
b	0,831	0,816
f_m	3 247 986,503 Hz	2 076 121,279 Hz
c_∞	$4{,}666\cdot10^{-8}$ F/km	$4{,}581\cdot10^{-8}$ F/km
c_0	$4{,}239\cdot10^{-9}$ F/km	$7{,}941\cdot10^{-11}$ F/km
c_q	$-2{,}344\cdot10^{-5}$	$-0{,}239\cdot10^{-5}$
g_e	0,459	0,482
g_0	$5{,}169\cdot10^{-5}$ S/km	$3{,}472\cdot10^{-5}$ S/km

6. Conclusion

This article describes methods for measuring of the characteristic impedance |Z0| using internal and external bridge, combined with automatic measurement facility controlled by computer in frequency range from 0,1 MHz to 100 MHz with the possibility to use up to 300 MHz. Facility can be used for attenuation, reflection and crosstalk measurement of twisted pairs. Such measuring facility can provide a measured waveforms and subsequent mathematical analysis. The measured values of secondary parameters UTP CAT5e cable will be used to refine the existing secondary parameters in the program used to evaluate the fundamental characteristics of metallic cables on Matlab server [3] and xDSL transmission rate estimation.

Acknowledgements

This work has been supported by the grant MSM6840770014.

References

[1] RAUSCHMAYER, D. J. *ADSL/VDSL Principles: A Practical and Precise Study of Asymmetric Digital Subscriber Lines and Very High Speed Digital Subscriber Lines*. Indianapolis, USA: Macmillan Technical Publishing, 1999. pp. 318. ISBN 978-1578700158.

[2] VODRÁŽKA, J.; SÝKORA, J.: New Methods for Subscriber Line Modeling. In *11th International Conference on Research in Telecommunication Technology* (RTT 2009). CTU FEE Prague 2009. pp. 67-70. ISBN 978-80-01-04410-0.

[3] VODRÁŽKA, J.; JAREŠ. P. *xDSL simulator*. Matlab server [online]. 2009. Available at WWW: <http://matlab.feld.cvut.cz/en/view.php?cisloclanku=2006110001>.

[4] LAFATA, P.; VODRÁŽKA, J. Modeling of Transmission Functions and Crosstalk in Metallic Cables for Implementation of MIMO Concept. *Radioengineering*. 2009, Vol. 18, No. 4, pp. 491-496. ISSN 1210-2512.

[5] VODRAZKA, J.; HRAD, J. *Modeling of a Subscriber Line with Inhomogeneity*. In *32nd International Conference on Telecommunications and Signal Processing* (TSP 2009), Dunakiliti, Hungary 2009. pp. 79-83. ISBN 978-963-06-7716-5.

About Authors

Lukas CEPA was born in Sumperk, Czech Republic in 1984. He received his M.Sc. degree in electrical engineering from CTU in Prague in 2009. His research interests include Internet Protocol version 6 (IPv6) and EMC of textile materials.

Milos KOZAK received the M.Sc. degree in electrical engineering from the Czech Technical University, Prague, in 2009. He participates in teaching of course of optical communication systems. His research interest is on the application of high-speed optical transmission systems in a data network. Currently he is actively involved in some of the projects on high speed optical modulations and Communication of Czech Antarctic Station.

Jiri VODRAZKA was born in Prague, Czech Republic in 1966. He joined the Department of Telecommunication Engineering, FEE, CTU in Prague in 1996 as a research assistant and received his Ph.D. degree in electrical engineering in 2001. He has been the head of the Transmission Media and Systems scientific group since 2005 and became Associate Professor in 2008. He participates in numerous projects in cooperation with external bodies. Currently he acts also as vice-head of the Department.

to Contents, page 1

SECURITY ANALYSIS SYSTEM TO DETECT THREATS ON A SIP VOIP INFRASCTRUCTURE ELEMENTS

Filip REZAC[1], *Miroslav VOZNAK*[1], *Karel TOMALA*[1], *Jan ROZHON*[1], *Jiri VYCHODIL*[1]

[1]Department of Telecommunications, Faculty of Electrical Engineering and Computer Science, VSB – Technical University of Ostrava, 17. listopadu 15, 708 33 Ostrava, Czech Republic

filip.rezac@vsb.cz, miroslav.voznak@vsb.cz, karel.tomala@vsb.cz, jan.rozhon@vsb.cz, jiri.vychodil@vsb.cz

Abstract. *SIP PBX is definitely the alpha and omega of any IP telephony infrastructure and frequently also provides other services than those related to VoIP traffic. These exchanges are, however, very often the target of attacks by external actors. The article describes a system that was developed on VSB-TU Ostrava as a testing tool to verify if the target VoIP PBX is adequately secured and protected against any real threats. The system tests the SIP element for several usually occurring attacks and it compiles evaluation of its overall security on the basis of successfull or unsuccessfull penetrations. The article describes the applications and algorithms that are used by system and the conclusion consists recommendations and guidelines to ensure effective protection against VoIP PBX threats. The system is designed as an open-source web application, thus allowing independent access and is fully extensible to other test modules.*

Keywords

SIP server, Safety Test System, Flood Attack, scanning, monitoring, SPIT, countermeasures, data manipulation.

1. Introduction

Systems designed to test and monitor networks or other components are quite wide-spread these days. Examples of the principle ones are Nessus [1], Retina [2], Snort [3] and other. The majority of these systems allows for testing the whole network infrastructures and protocols used for communication between components. None of these solutions, however, enables a complex testing of VoIP infrastructure and SIP servers which are the key and most vulnerable component of the network. The system we developed, under a working title SPT (SIP Penetration Testing), was designed as a penetration tests simulator for SIP servers. Based on the analysis of intersections, the person who initiated the testing ("the tester") receives feedback in the form of test results, as well as recommendations on how to mitigate potential security risks that were discovered. The advantage of this solution is that the system simulates real attacks from the external network, i.e. the system does not need to be placed in the same network as the target component DUT (Device under Test). This is frequently one of prerequisites to be able to use other testing tools. The SPT system was implemented as a web application accessible through a standard web browser and therefore independent on the operation system's platform. Authentication will be done using the SSO (Single Sign-On) service - Shibboleth [4]. This should also prevent the system being used for other than testing purposes. Once signed in, the tester enters the required data into a web form and chooses tests to be run. The output of the application once the tests have been completed is an e-mail report to the tester.

This paper contains the results of the tests; and in case some penetrations were successful it also contains recommendations and measures to mitigate such attacks in the future. Figure 1 illustrates the concept of the SPT system. The following chapter describes individual testing methods in detail, their implementation, algorithms used and the impact on the target SIP server.

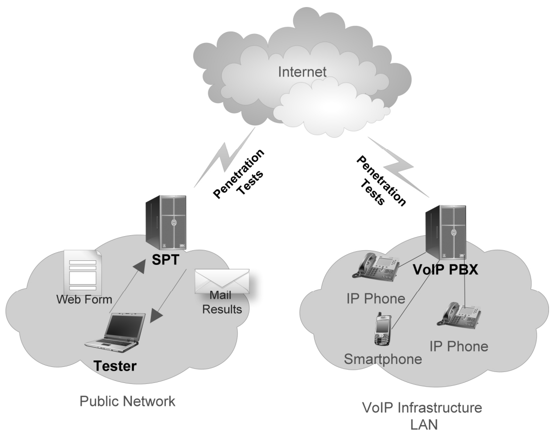


Fig. 1: SIP penetration tests system scheme.

to Contents, page 1

2. Platforms, Algorithms and their Time Evaluation

Although the system is primarily designed for penetration tests on SIP servers, in reality it can perform full-scale attacks on a particular component and provide feedback on it to the tester. Thus, it is necessary to ensure that the developed system cannot be abused by a third party. The system was designed as a LAMP (Linux, Apache, MySQL, PHP) server [8] and its complete administration including the installation is carried out via a web interface. For reasons stated above, the system will only be accessible to authorised persons once they pass through the authentication. First the tester fills in the IP address or domain name of the central SIP server and the email address to which the test results will be sent. Using checkboxes, the tester may define the range of the modules offered for testing. Individual modules are described below in detail.

2.1. Scanning and Monitoring Module

In order to be able to carry out an efficient and precise attack on a SIP server, the potential attacker needs to find as much information as possible about a particular component. This is why we first developed a Scanning and Monitoring ("S&M") module for the SPT system, which is used to test the security of the server against attacks aimed at obtaining information by means of common and available tools (Fig. 2).

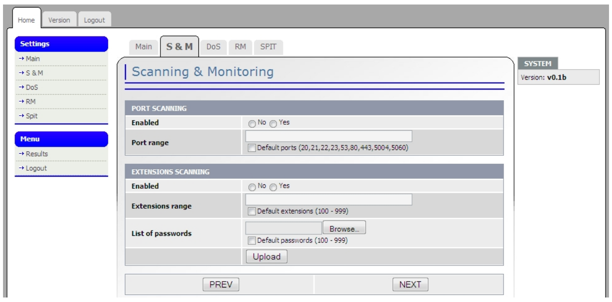


Fig. 2: SPT system – S&M module.

These tools include for instance Nmap [9] or ever more popular SIPvicious [10]. SPT system also uses these testing tools. By means of these tools, it is possible to obtain a list of listening ports or a list of user accounts created from the unsecure server. Where the server is not secured sufficiently, they can obtain even the most important - passwords to individual accounts. If the tester ticks the test to be carried out, the Nmap application is used first to establish open ports. Given the time requirements of the T_n [s] test, the testing is by default restricted only to several most frequently used ports. Using the web form, the tester can set the range of the tested ports. However the total time set for testing using Nmap is 1800 s (30 minutes). The list of available ports is subsequently included in the assessment report together with recommendations on how to minimize such ports' scanning. Another test which the SPT system can carry out aims at establishing whether SIP server's security allows for obtaining a list of user accounts. For this purpose, SIPvicious is used. By sending out OPTION and ACK requests, the application detects what accounts are defined on the SIP server. By default, the system tries the 100-999 range of accounts. Again, the tester may define own range of tested numbers E_{nr} or import a text file containing strings of alpha- numeric characters or words E_{dr}. Time required to check and create a list of T_e [s] accounts can be expressed by Eq. (1) where c = 0,02603 is a time constant obtained by repetitive measurements on a sample of 1000 potential accounts on different target SIP servers [7].

$$T_e = (E_{nr} + E_{dr}) \cdot c\,. \tag{1}$$

Number of valid accounts E_{valid} is derived from Eq. (2) where $E_{invalid}$ is the number of accounts that have been reviewed by the system but not defined on the SIP server.

$$E_{valid} = (E_{nr} + E_{dr}) - E_{invalid}\,. \tag{2}$$

Once the system has tested security of the SIP server against detecting accounts, possibility to detect passwords for individual accounts is tested. Again, this testing is carried out by *SIPvicious*. Using a pre-defined range of possible numeric passwords P_{nr} or an imported text file with alpha-numeric characters or words P_{dr}, it obtains a list of passwords for individual accounts. Time requirements on this test are expressed by the following Eq. (3).

$$T_p = \left[E_{valid} \cdot (P_{nr} + P_{dr})\right] \cdot c\,, \tag{3}$$

$$T_{sm} = T_e + T_p + T_n\,. \tag{4}$$

Now we can determine the estimated time required to carry out the complete S&M test T_{sm} (4). Using the module, we can verify whether the target SIP server is sufficiently secured against such scanning and monitoring attacks.

2.2. Denial of Service Module

One of the most frequently occurring attacks is DoS (Denial of Service). In reality, it consists of several attacks with the same characteristic feature – to lock up or restrict the availability of the attacked service so that it does not function properly. Several types of DoSs [11] can be used to achieve this; our system tests the SIP server using the most frequently used one, Flood DoS. The principle of the attack is to send a large volume of adjusted or otherwise deformed packets to the target component so that it is unable to provide its core services. As a result of the attack, CPU load increases and most of the available bandwidth is consumed, resulting in the SIP

to Contents, page 1

server being unable to service regular calls, or only a minimum amount of them. To generate Flood DoS, the SPT system uses two applications: udpflood [12] and inviteflood [12]. When using udpflood, the system generates UDP packets of 1400 bytes which are directed at SIP default port 5060 of the target SIP server. The tester defines the number of generated packets and the system tests whether the packets arrived at the SIP server and whether they cause some restriction of the service availability (Fig. 3). Since we know the packet's size and therefore also the size of the Ethernet framework Fs_{udp}, we can, based on the number of generated packets P_n and the bandwidth provided B_w, determine time T_{udp} [s] required to carry out the test (5).

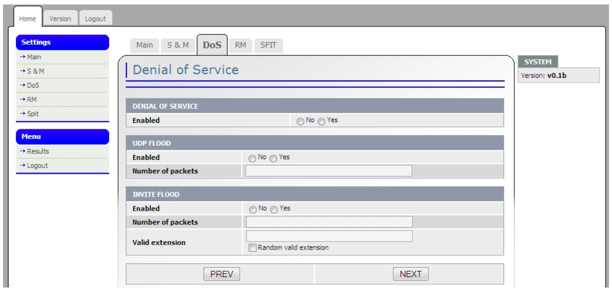


Fig. 3: SPT system - DoS module.

$$T_{udp} = \left(Fs_{udp} \cdot P_n\right) / B_w . \tag{5}$$

Table 1 provides an overview of time required for different numbers of generated packets P_n and different bandwidth B_w. When the other application, *inviteflood*, is used for testing, the system generates INVITE requests at the SIP server which are directed at an existing account. This method is very successful as most of today's SIP servers require an authentication for INVITE requests. As the INVITE requests generated by our system do not contain any authentication string, the SIP server returns SIP answer 407 Proxy Authentication Required. With the large volume of incoming requests, the load of SIP server's CPU increases. The tester can set the value of a valid account in the system manually, or it can be randomly selected from the previously obtained list of valid accounts E_{valid}. As in the previous case, we can, based on the number of generated packets P_n and the bandwidth provided B_w, determine time T_{invite} [s] required to carry out the test (6).

$$T_{invite} = \left(Fs_{invite} \cdot P_n\right) / B_w . \tag{6}$$

Figure 4 illustrates the impact of the change in bandwidth on CPU load when simulating an *udpflood* attack. The chart also clearly shows resistance of the two popular open-source SIP servers, Asterisk PBX [5] and OpenSIPS [6], to UDP Flood DoS attacks. Both centrals have been installed on the same HW of Dell PowerEdge R510 server to eliminate any potential difference in computational performance. To change bandwidths, we used HW emulator of the Simena networks. CPU load on individual centrals was measured by means of *dstat* [13].

The chart shows that OpenSIPS is many times more resistant to UDP DoS attacks than Asterisk. Total time required to carry out DoS tests T_{dos} is determined as follows (7).

Tab.1: Udpflood attack time duration with different bandwidth and number of generated packets.

Number of packets - Pn	Bandwidth [Mbit·s^{-1}] and the attack time Tudp [s]			
	10	**25**	**50**	**100**
100 000	113,12	45,25	22,63	11,31
200 000	226,24	90,50	45,26	22,62
300 000	339,36	135,75	67,89	33,93
400 000	452,48	181	90,52	45,24
500 000	565,60	226,25	113,15	56,55
600 000	678,72	271,5	135,78	67,86
700 000	791,84	316,75	158,41	79,17
800 000	904,96	362	181,04	90,48
900 000	1018,08	407,25	203,67	101,79
1 000 000	1131,20	452,5	226,3	113,1

$$T_{dos} = T_{udp} + T_{invite} . \tag{7}$$

Results and success rate of DoS tests carried out are included in the report for the tester.

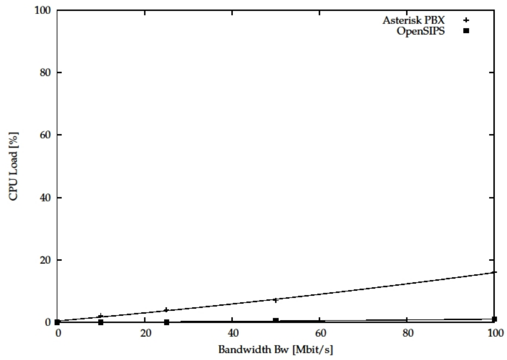


Fig. 4: Impact of change in bandwidth on CPU load in case of *udpflood* attack.

2.3. Registration Manipulation Module

Once the potential perpetrator obtains information about existing accounts, he can manipulate these accounts quite easily. The SPT system we developed can also test SIP servers' security, i.e. measures against manipulating the registration, see Fig. 5.

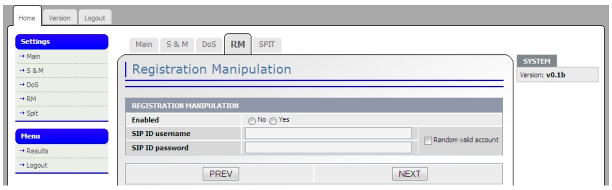


Fig. 5: SPT system - RM module.

to Contents, page 1

To carry out this test, the system uses *reghijacker* [12] which substitutes the legitimate account registration with a fake, non-existing one. This type of attack can easily be expanded to a so called MITM, Man-in-the-Middle [11]. In this attack, a non-existent user is substituted by a valid SIP registration and all incoming signaling and media to the legitimate registration will be re-directed to the newly created registration. In this case, the tester needs to define the value of the SIP account which is to be stolen in the system and where authentication of REGISTER request is allowed, also a password to this account. Where the tester fails to define these values, the system automatically assigns an account and its password from the list created while scanning and monitoring the central. Time required to carry out the test T_{rm} is insignificant compared to operational times of other modules.

2.4. SPIT Module

Today, one of the most popular attacks on the Internet is spam. It is estimated that spams account for 80 – 90 % of total attacks on the Internet. Security experts predict that Spam over Internet Telephony (SPIT) will be a major threat in the future. The level of annoyance is even greater than with classical spam. Our team had developed SPITFILE [14] which served as a testing tool while developing security against such type of attacks. The SPT system uses the core of this application, together with *Sipp* [14], to simulate a SPIT attack on the target SIP server (Fig. 6). In the form, the tester defines the value of a valid SIP account – the called party to which the SPIT call will be directed and then the value and password to a valid SIP account – the caller through which the call will be initiated. Where the tester fails to define these values, the system automatically assigns an account and an appropriate password from the list created while scanning and monitoring the central.

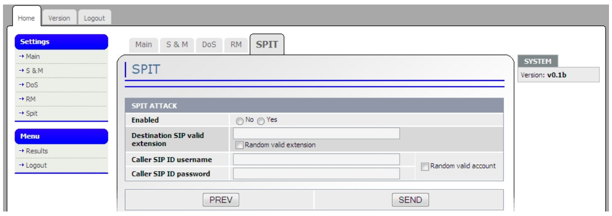


Fig. 6: SPT system - SPIT module.

If the attack was successful, a SIP call is initiated from the caller's account, and the end device with the registered account of the called party starts ringing. Once the call is answered, a pre-recorded message is played and the call terminated. Time required to carry out the test T_{spit} is determined by the length of the pre-recorded message. The final report on penetration tests which the tester receives via e-mail, will, besides information on all previous tests, also contain an analysis and success rate of the SPIT module's test.

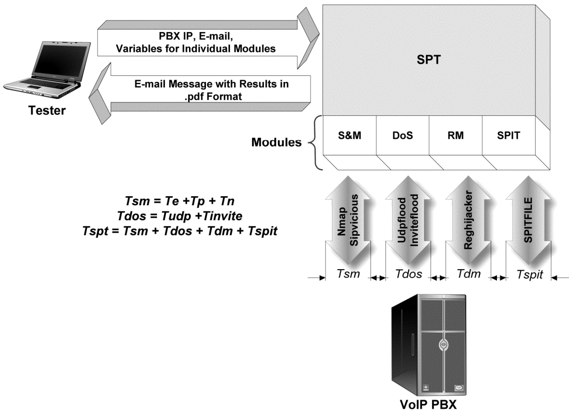


Fig. 7: Division of the SPT system into individual modules.

Figure 7 illustrates the division of the SPT system into individual modules and shows time intervals necessary to carry out individual tests in respective modules. Time requirements of the whole SPT system can be expressed by Eq. (8). Its value depends on many factors and can radically change in accordance with the type of tests requested by the tester. Its value is for reference only.

$$T_{spt} = T_{sm} + T_{dos} + T_{rm} + T_{spit} . \tag{8}$$

3. Platforms, Algorithms and their Time Evaluation

Although the SPT system is still in the phase of intensive testing and development, basic operational tests of all available modules were carried out. Each test is accompanied by a short description of countermeasure's principles and methods [12] which should limit or completely mitigate potential security gaps that were revealed during SIP servers' testing.

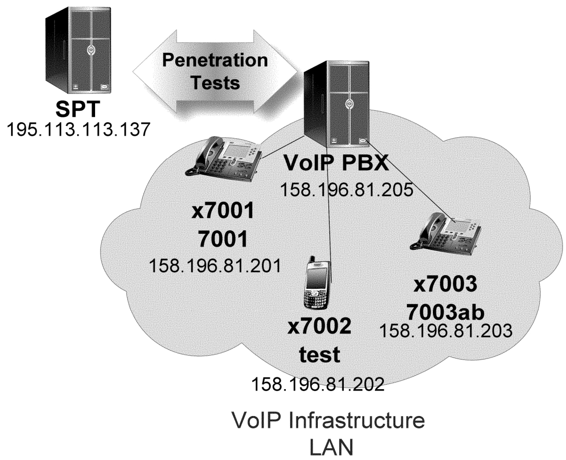


Fig. 8: SIP penetration tests system testbed.

to Contents, page 1

Figure 8 describes the basic testing topology. The system is denoted as SPT and Asterisk as VoIP PBX. Asterisk was installed at Dell PowerEdge R510 server.

3.1. Scanning and Monitoring Module Testing and Countermeasures

The first step was to enter IP address of SIP server and e-mail address to which the report will be sent. Next, the S&M module and subsequently *Nmap* and *SIPvicious* applications were launched. Values for *Nmap* were set by default, value of E_{nr} for *SIPvicious* was set to range between 1000-9999. The device found all three registered accounts E_{valid} 7001-7003 and listed open TCP and UDP ports at Asterisk. Once P_{nr} was set to 7001-7003 and a text file P_{dr} containing test and 7003ab string, the test to obtain passwords to individual accounts was also successful (Fig. 10a). Total time incurred on testing module is $T_{sm} \cong 235$ s. Part of the report that is delivered to a tester's email is shown on Fig. 9. If we had to protect and prevent SIP server from scanning and monitoring, then an implementation of firewall is the effective solution or an intrusion detection system that is able to distinguish scanning and monitoring. The next effective solution is to divide the network logical infrastructure into VLANs and decompose the provided services into more physical servers (TFTP, HTTP servers). The prevention of accounts and passwords detection is difficult, moreover, the tools for detection apply the standard SIP methods and is not trivial to distinguish legitimate behaviour from an attack. In this case, it is recommended to divide the infrastructure into individual VLANs so that the detection for intruder is as difficult as possible.

```
Result of Scanning and Monitoring test
==================================================================
Tested Device: 158.196.[illegible]:5060
Agent On Device: Asterisk PBX
Port Scan Results:
[20/tcp] => filtered
[21/tcp] => filtered
[22/tcp] => filtered
[23/tcp] => filtered
[53/tcp] => filtered
[80/tcp] => filtered
[443/tcp] => open
[5004/tcp] => filtered
[5060/tcp] => closed
[20/udp] => open|filtered
[21/udp] => open|filtered
[22/udp] => open|filtered
[23/udp] => open|filtered
[53/udp] => open|filtered
[80/udp] => open|filtered
[443/udp] => open|filtered
[5004/udp] => open|filtered
[5060/udp] => open|filtered

Port 443/tcpis open!!

System found extensions:
[[illegible]01] => reqauth
[[illegible]02] => reqauth
[[illegible]03] => reqauth


System does not found any valid password for extensions.
```

Fig. 9: Example of the SaM module test results that are sent to Tester's email.

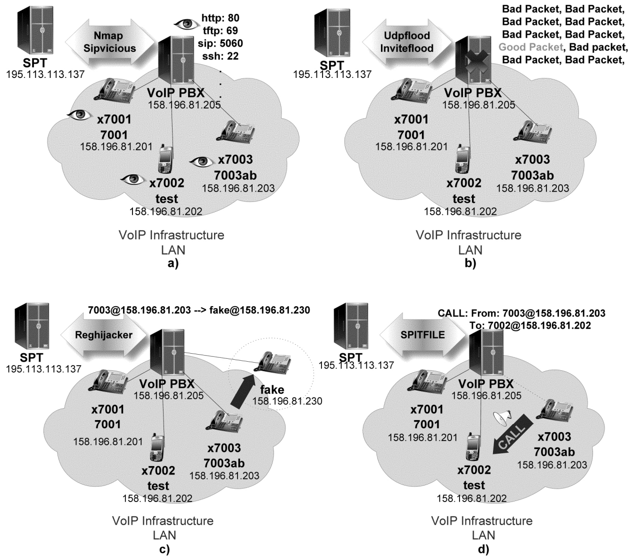


Fig. 10: SIP penetration tests System testbed. a) scanning and monitoring tests, b) DoS tests, c) registration manipulation tests, d) SPIT tests.

3.2. DoS Module Testing and Countermeasures

Using *udpflood*, the tester sent 500000 UDP packets directly to port 5060. Bandwidth was set to 100 Mbit/s, Asterisk processed 90 % calls. Once the test was completed, Asterisk recovered to a full operation mode. To be able to compare, we substituted Asterisk by OpenSIPS in this test. Call processing under the same attack was entirely error-free. When testing using *inviteflood* on the valid account 7001, we found out that this attack is much more destructive in terms of computational power. As early as at 100000 INVITE request when $T_{invite} \cong 9$ s, CPU load for both Asterisk and OpenSIPS reached 100 % and failed to process a single incoming or outgoing call. Once the test was completed, both centrals recovered to a full operation mode (Fig. 10b).

The way how to recognize whether a DoS attack is successful or not is given below. Before the test starts, SPT uses the ICMP protocol to measure the average response time from the exchange - T_{avg}. During testing (sending flood packets) the average response - T_{dosavg} is again tested in parallel using ICMP. In case that the T_{dosavg} is 150 times greater than T_{avg}, the test is marked as successful. A necessary condition for DoS test is to support an ICMP protocol. The principle of detection is shown in Fig. 11.

to Contents, page 1

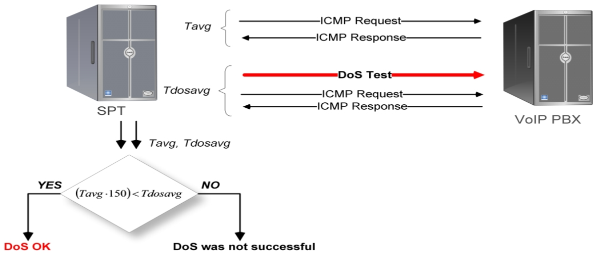


Fig. 11: Principle of DoS test detection.

The possibilities on how to protect from Flood DoS attacks, are the following: to divide the network infrastructure into separate VLANs, to have in use solely TLS, to implement L2 network elements with DoS detection or to apply SIP firewall that can detect DoS attacks and minimize their impact.

3.3. Registration Manipulation Module Testing and Countermeasures

When testing possibility for registration manipulation, we entered values of account 7003 and its password 7003ab manually into the SPT system. Once the test was completed, we established whether the attack was successful using the following process. SPT sends a SIP INVITE message to the test machine. Once a message is sent, the system waits for the type of response that comes. Based on this response SPT assess whether the test was (attack), successful or not. When system receives **180 Ringing** response the tested attack has not been successful, if a **503 Service Unavailable** response arrives, the test was implemented successfully (see Fig. 12).

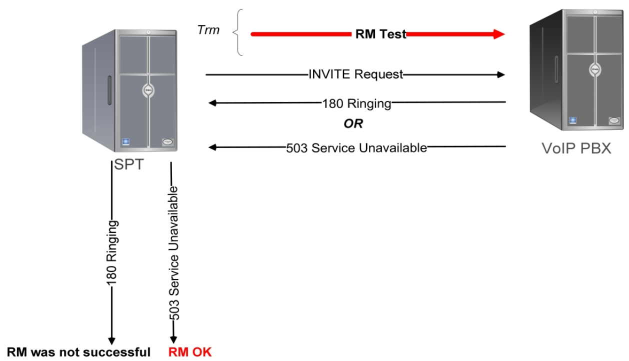


Fig. 12: Principle of RM test detection.

The aim of the attack was to de-register account 7003 and to direct all incoming calls to a fake account which does not exist. Thus, calls were terminated as unconnected. The call to 7003 was not put through, see Fig. 10c. The TCP protocol is recommended at transport level to prevent a registration hijacking because the manipulation with TCP requires higher complexity. Next option, how to minimize this threat, is to use REGISTER message authentication. We could decrease the registration interval, as well, it is quite simple but effective.

3.4. SPIT Module Testing and Countermeasures

As stated above, we used SPITFILE application, developed by authors of the paper, to test the PBX vulnerability against SPIT attacks. The tester manually inserts value of a valid account 7002 on which a SPIT attack was to be initiated, as well as the value of a valid account 7003 and password to it (7003ab) which was supposed to initiate the SPIT call. Once the test was launched, SPITFILE registered on the participant 7003 and then started to generate a call to account 7002. The end device registered on 7002 began ringing, and once the call was answered, a recording with an advertisement was played (Fig. 10d). The system detects the success of SPIT attacks as follows: during test, SPT monitors the responses of SIP VoIP PBX and based on the them, system defines whether the SPIT test was successful or not. System performs 5 consecutive SPIT calls, and all INVITE messages must be answered with **180 Ringing** response. If SPT receives any other response, SPIT test failed. Detection process can be seen in Fig. 13.

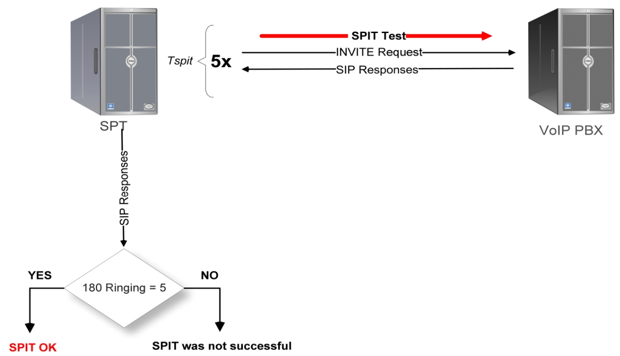


Fig. 13: Detection of SPIT test.

A few methods exist on how to restrict the SPIT propagation, which are more or less efficient, but their combination bring quite strong protection against the type of attack. Such methods include various forms of lists or voice testers based on CAPTCHA which are particularly effective against caller bots. Authors developed own solution ANTISPIT [14] that exploits the specific human behavior and automatically modifies the Blacklist table without participation of called party, the approach is based on the statistical Blacklist.

to Contents, page 1

4. Conclusion and Future Work

The aim of the authors was to develop a tool to carry out penetration tests on SIP servers. The system that was designed and implemented consists of several modules that are able to generate selected types of attacks which the authors deem most popular. The system then analyses to what extent is the target component secured, drafts assessments containing tests' results and proposes factual recommendations to ensure security against the threat concerned. The assessment report is sent as a text document to an e-mail. The system is currently under intensive testing. It is planned that in the future, it will be extended to include other testing modules and functions such as for instance testing of the whole VoIP infrastructure and heavy testing of individual components.

Acknowledgment

This work has been supported by the Ministry of Education of the Czech Republic within the project LM2010005.

References

[1] ROGERS, R. *Nessus Network Auditing. Syngress.* 2. Burlington: Elsevier, 2008. pp. 433. ISBN 978-1-59749-208-9.

[2] CHOCHELINSKI, R.; BARONAK, I. Private Telecommunication Network Based on NGN. In *32nd International Conference on Telecommunications and Signal Processing*, 2009, Dunakiliti, HUNGARY, pp. 162-167. ISBN 978-963-06-7716-5.

[3] BATES, J.; GALLON, C.; BOCCI, M.; WALKER, S; TAYLOR, T. *Converged Multimedia Networks*. Wiley, 2006. pp. 364. ISBN 978-0-470-02553-6.

[4] VOZNAK, M. *Voice over IP*. VSB-Technical University of Ostrava: College Textbook, 1st. ed., Ostrava, 2008.

[5] WINTERMEZER, S.; BOSCH, S. *Practical Asterisk 1.4 and 1.6: From Beginner to Expert*. Addison-Wesley Professional; 1 edition, 2009. ISBN 978-0-321-52566-6.

[6] GONCALVEZ, F. *Building Telephony Systems with OpenSIPS 1.6*. Packt Publishing, 2010. pp. 274, ISBN 978-1849510745.

[7] SISALEM, D.; FLOROIU, J.; KUTHAN, J.; ABEND, U.; SCHULZRINNE, H. *SIP Security*. Wiley, 2009. pp. 350. ISBN 978-0-470-51636-2.

[8] LEE, J.; WARE, B. *Open Source Development with LAMP: Using Linux, Apache, MySQL, Perl, and PHP*. Addison-Wesley Professional, 2002. pp. 496. ISBN 0-785-342-77061-2.

[9] LYON, G. F. *Nmap Network Scanning: The Official Nmap Project Guide to Network Discovery and Security Scanning*. Nmap Project, 2009.

[10] VOZNAK, M.; ROZHON, J. SIP Infrastructure Performance Testing. In *9th International Conference on Telecommunications an Informatics*, Catania, Italy, 2010, pp. 153-158. ISBN 978-954-92600-2-1.

[11] REZAC, F.; VOZNAK, M.; RUZICKA, J. Security Risks in IP Telephony. In *CESNET Conference 2008*, 2008, Prague, pp. 31-38. ISBN 978-80-904173-0-4.

[12] ENDLER, C.; COLLIER, M. *Hacking Exposed VoIP: VoIP Security Secrets and Solutions*. McGraw-Hill Companies, 2007. pp. 539. ISBN 978-0-07-226364-0.

[13] PRAVDA, I.; VODRAZKA, J. Voice quality planning for NGN including mobile networks. In *12th International Conference on Personal Wireless Communications (PWC 2007)*, 2007, Prague. pp. 680. ISBN 978-0-387-74158-1.

[14] VOZNAK, M.; REZAC, F. The implementation of SPAM over Internet telephony and defence against this attack. In *TSP 2009: 32nd International Conference on Telecommunications and Signal Processing*, Dunakiliti, HUNGARY, Aug 26-27, 2009, pp. 200-203. ISBN 978-963-06-7716-5.

About Authors

Filip REZAC was born in 1985. In 2007, received a Bachelor title in VSB-Technical University of Ostrava, Faculty of Electronics and Computer Science, Department of Informatics. Two years later he received the M.Sc. title focused on mobile technology in the same workplace. Currently in the doctoral study he focuses on Voice over IP technology, Network Security and Call Quality in VoIP.

Miroslav VOZNAK was born in 1971. He holds position as an associate professor with Department of Telecommunications, VSB-Technical University of Ostrava, Czech Republic. He received his M.S. and Ph.D. degrees in telecommunications, dissertation thesis "Voice traffic optimization with regard to speech quality in network with VoIP technology" from the Technical University of Ostrava, in 1995 and 2002, respectively. Topics of his research interests are the next generation network, IP telephony, speech quality and network security.

Karel TOMALA received a Bachelor title in VSB-Technical University of Ostrava, Faculty of Electronics and Computer Science, Department of Electronics and Communications in 2007. Two years later he received the M.Sc. title on the Department of Telecommunications. Currently in the doctoral study he focuses on Voice over IP technology, High Availability on Telecommunication Networks and Embedded Solutions.

Jan ROZHON received his M.Sc. degree in telecommunications from VSB – Technical University of Ostrava, Czech Republic, in 2010 and he continues in studying Ph.D. degree at the same university. His research is focused on performance testing of NGN. In 2010, he received rector's appreciation for his diplomathesis.

to Contents, page 1

Jiri VYCHODIL was born in 1984 in Olomouc. In 2007, he received bachelor degree at VSB-Technical University of Ostrava, Faculty of Electronics and Computer Science, Department of Telecommunications. In 2009 he received M.Sc. degree at the same department. Now, he is a Ph.D. student. His research is focused on IP telephony and computer networks.

to Contents, page 1

E-MODEL MOS ESTIMATE IMPROVEMENT THROUGH JITTER BUFFER PACKET LOSS MODELLING

Adrian KOVAC[1], *Michal HALAS*[1], *Milos ORGON*[1], *Miroslav VOZNAK*[2]

[1]Department of Telecommunications, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, Ilkovicova 3, 812 19 Bratislava, Slovakia
[2]Department of Telecommunications, Faculty of Electrical Engineering and Computer Science, VSB-Technical University of Ostrava, 17.listopadu 15, 803 00 Ostrava, Czech Republic

fotoadrian@gmail.com, halas@ut.fei.stuba.sk, orgon@ut.fei.stuba.sk, miroslav.voznak@vsb.cz

Abstract. *Proposed article analyses dependence of MOS as a voice call quality (QoS) measure estimated through ITU-T E-model under real network conditions with jitter. In this paper, a method of jitter effect is proposed. Jitter as voice packet time uncertainty appears as increased packet loss caused by jitter memory buffer under- or overflow. Jitter buffer behaviour at receiver's side is modelled as Pareto/D/1/K system with Pareto-distributed packet interarrival times and its performance is experimentally evaluated by using statistic tools. Jitter buffer stochastic model is then incorporated into E-model in an additive manner accounting for network jitter effects via excess packet loss complementing measured network packet loss. Proposed modification of E-model input parameter adds two degrees of freedom in modelling: network jitter and jitter buffer size.*

Keywords

VoIP, E-model, MOS, QoS, packet loss, jitter buffer, loss estimate, Generalized Pareto, packet interarrival, Pareto/D/1/K.

1. Introduction

Mean opinion score - MOS - represents user satisfaction with overall listening and conversational quality on five grade scale from 5 (best) to 1 (worst). MOS can be estimated by either subjective methods based on user listening tests or by objective methods based on measured network parameters such as delay, packet loss and jitter. E-model defined by ITU-T G.107 [1] is widely accepted objective method used for estimation of VoIP call quality. E-model uses set of selected input parameters to calculate intermediate variable – R factor, which is consecutively converted to MOS value using (3). Input parameters contribute to final estimate of quality in additive manner as expressed in Eq. (1).

$$R = R_o - I_s - I_d - I_{e,eff} + A, \qquad (1)$$

where R_o represents the basic SNR, circuit and room noise; I_s represents all impairments related to voice recording such as quantization distortion, low voice volume; I_d covers degradations caused by delay of audio signal including side-tone echo; $I_{e,eff}$ impairment factor presents all degradations caused by packet network transmission path, including end-to-end delay, packet loss and codec PLC masking capabilities; A is an advantage factor of particular technology.

Tab.1: Impairment factors for selected codecs [5].

Audio Codec	Codec bitrate	Impairment Factor	
		I_e	B_{pl}
G.711 w/o PLC	64 kb/s	0	10
G.711 with PLC	64 kb/s	0	34
G.723.1	5,3 kb/s	19	24
G.723.1	6,3 kb/s	15	20
G.726	16 kb/s	40	69
G.726	24 kb/s	25	38
G.726	32 kb/s	12	24
G.726	40 kb/s	7	24
G.728	16 kb/s	16	27
G.729	8 kb/s	10	18
G.729 A	8 kb/s	11	17
GSM FR 6.10	13,2 kb/s	26	43

We focus at $I_{e,eff}$ parameter, which is calculated as in Eq. (2):

$$I_{e,eff} = I_e + (95 - I_e) \cdot \frac{P_{pl}}{P_{pl} + B_{pl}}, \qquad (2)$$

to Contents, page 1

where I_e represents impairment factor given by codec compression and voice reproduction capabilities, B_{pl} is codec robustness, which describes how immune is particular codec against random losses and what are its PLC masking qualities. These values are given for narrowband codecs in ITU-T G.133 appendix [5]. Table 1 gives a summary of I_e and B_{pl} parameters for selected codecs. P_{pl} parameter represents measured network packet loss in %.

2. Relations between Network Delay, Jitter and Packet Loss

An easiest way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.

2.1. Properties of Network Delay

Voice packets being sent from VoIP device – IP phone – can be considered as a regular flow with constant transmit intervals and transmit duration. VoIP packets after transport network traversal have their regular time spacing disrupted in irregular way. Internet traffic arrival times and delay can be successfully statistically modelled by long-tailed Generalised Pareto distribution (GPD) [7], [9], [10] which we use to describe incoming VoIP packet stream from transport network to IP phone and refer to it like "Pareto distribution" further in this paper. Network delay distribution of received packets is depicted in Fig. 1.

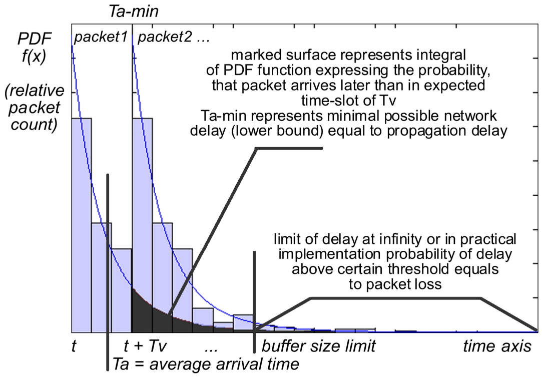


Fig. 1: Packet transmission delay distribution.

Real-time change of network operational and performance parameters cause variations in delay. Differences between consecutive packet arrivals are not constant and arrival times oscillate between minimal possible network delay *Ta-min* and infinite delivery time (lost packet). Mean value of the described process exists and is called and measured as an End-to-End network delay *Ta* (one of the input parameters for E-model).

2.2. Network Delay Statistics

Real packet path usually consists of a mixture of different networks with different devices and technologies. Each device adds certain degree of uncertainty in packet processing and transmitting time. Overall network delay statistics is a result of all partial statistics at each device. Aforementioned Pareto distribution is well suited to describe time as a variable, which has lower bound, no upper bound and finite mean value. It is better suited to describe internet packet flow than Weibull distribution as it yields lower MSE in practical measurements of real traffic [7], [9], and [10]. Pareto has different shape than exponential distribution with steeper slope and less occurrences in higher values. Probability density function of Pareto (PDF) is given by Eq. (3) and cumulative distribution function (CDF) by Eq. (4).

$$f_{(\xi,\mu,\sigma)}(x)=\frac{1}{\sigma}\left(1+\frac{\xi(x-\mu)}{\sigma}\right)^{\left(-\frac{1}{\xi}-1\right)}, \qquad (3)$$

$$F_{(\xi,\mu,\sigma)}(x)=1-\left(1+\frac{\xi(x-\mu)}{\sigma}\right)^{\left(-\frac{1}{\xi}\right)}, \qquad (4)$$

where σ = std. deviation, ξ = shape parameter, μ = location parameter (minimal value of random variable with Pareto distribution). μ represents offset of Pareto distribution from zero on time axis and represents minimal network delay *Ta_min* also illustrated at Fig. 1. Shape of Pareto PDF for varying σ is depicted at Fig. 2. For Pareto PDF to converge faster than exponential, shape parameter must meet condition $\xi < 0$ and to get valid results from Eq. (3) and (4) $\mu \leq x \leq \mu - \sigma/\xi$.

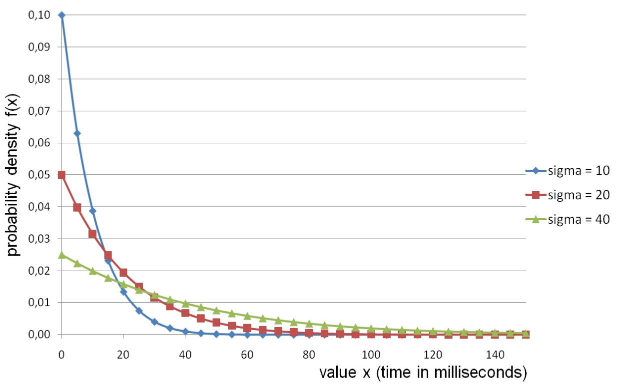


Fig. 2: Packet transmission delay distribution.

2.3. Network Jitter

Jitter J [ms] is calculated in real-time as floating average of differences between interarrival times called "Timestamps" of consecutively received packets contained in RTP protocol header. Calculation of *J* is given by Eq. (5). Each particular difference is given by

to Contents, page 1

Eq. (6), both according to RFC 1889, where R are timestamps when packet was received, S when packet was sent and indices i,j are consecutive packet numbers. Jitter value is transferred in RTCP protocol header as one of the QoS parameters. For correct measures of jitter correct synchronization of clocks in network is needed.

$$J = J + \left(\left|D_{(i-1,i)}\right| - J\right)/16\left[ms\right], \tag{5}$$

$$D_{(i,j)} = \left(R_j - R_i\right) - \left(S_j - S_i\right)\left[ms\right], \tag{6}$$

During VoIP communication, voice packets are generated in regular intervals with fixed period Tv assuming that no VAD is active and packetization interval is constant. Deviations of placing packets in correct timeslots on transport medium can occur when other traffic is present, e.g. when sending files through FTP and simultaneously making VoIP calls, traffic has to be interleaved. These issues are further amplified at active layer 2 and 3 network elements when routing, switching and traffic shaping occurs. Receiver expects packets at regular timeslots but to compensate disturbances in delivery time receiver has to use buffer memory called “jitter buffer” at the expense of increasing overall end-to-end delay at least by length of buffer in milliseconds / 2. Result of the process at receiver’s side is an irregular packet stream which can be described through Markov process as Pareto/D/1/K queue [7], [9] depicted in Fig. 3.

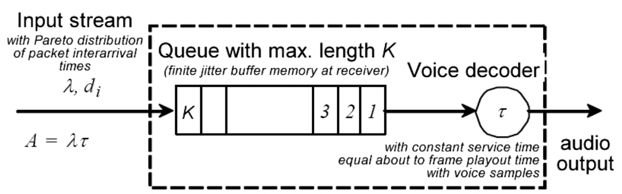


Fig. 3: Markov chain model of VoIP packet receiver with buffer.

3. Analysis of Jitter Effects

3.1. Effective Packet Loss

Packet Loss in packet data network can occur either in transport network or at user device. Transport network can suffer from congestion, load imbalance and also mixture of several transport technologies, which behave differently.

Jitter itself is not one of the input parameters for E-model calculation, but its influence is not negligible under real network conditions. Effects of jitter on packet loss appear as increased net total packet loss which is equal or greater than measured network packet loss expressed as P_{pl}. In this document we refer to “Modified E-model” as to model based on original ITU-T G.107 [2] recommendation with P_{pl} packet loss parameter in Eq. (2) substituted by proposed P_{plef}. Codec robustness factor B_{pl} in Eq. (2) expresses resilience of certain codec to random and burst packet loss and its masking capabilities through packet loss concealment (PLC) algorithm. Overall effective packet loss P_{plef} can be expressed as a complement to product of packet transmittance in network and in receiver jitter buffer as given by Eq. (7).

$$P_{plef} = 1 - \left(1 - P_{pl}\right)\cdot\left(1 - P_{jitter}\right) \quad \in \langle 0;1\rangle, \tag{7}$$

that can be rewritten to Eq. (8) as a jitter buffer loss:

$$P_{jitter} = \left(P_{plef} - P_{pl}\right)/\left(1 - P_{pl}\right) \quad \in \langle 0;1\rangle, \tag{8}$$

where P_{pl} is network packet loss as measured by receiving device and P_{jitter} is an additional packet loss occurring at finite jitter buffer memory when compensating time fluctuations of arriving voice packets. Packet transmittance is a complement to packet loss and is expressed by terms in brackets in Eq. (7).

3.2. Jitter Buffer Behaviour

We have used Pareto/D/1/K queue [7], [9] depicted in Fig. 3 to model jitter buffer behaviour. Markov chain Pareto/D/1/K behaves like M/D/1/K because memory-lessness of previous states of input stream is ensured. Zero correlation of Poisson input stream is maintained with Pareto distribution.

Figure 4 shows sample timelines of transmitter and receiver of VoIP packets, expected arrival timeslots as well as deviations from expected delivery time, for which there is a jitter buffer. Jitter buffer has threshold levels of how much ms of audio samples it has to receive before starting playback of stream, which is usually 50 % of its length and it equals to additional end-to-end delay added by buffer. If receiving buffer behaves like ordinary queue, packets that are severely influenced by jitter and delayed so much that they arrive later than packets sent after them, such exchange in packet order cannot be treated and late packet is treated as lost even when its timeslot has not passed (Fig. 4). Real jitter buffer can repair or reorder swapped packets in input stream by their RTP sequence number when their timeslot or time to play has not passed. This process significantly lowers additional packet loss in networks with high jitter comparable with packet length (jitter usually 30 ms and higher).

For calculation of P_{jitter} at jitter buffer with reordering capability in Eq. (8) we have to recalculate P_{jitter} in raw Markov chain-type buffer without reordering capability represented by PDF Pareto function $F_{wo}(x,\mu,\xi,\sigma)$ (transmittance of buffer without reordering capability) to F_{wr} (transmittance of buffer with packet reordering capability) according to Eq. (9). Formula (9) is based on autocorrelation function of packet loss P_{jitter} without reordering which equals to probability, that packet will not be in jitter buffer present at the moment of processing in its expected timeslot. Real measured values

to Contents, page 1

of F_{wo}, F_{wr} as well as their complements – jitter buffer packet losses P_{wo} and P_{wr} are present in Tab. 2 and 3. These values were used to find the best fitting shape parameter with minimal MSE error of packet losses on buffer with reordering ability.

$$P_{jitter} = F_{wr} = \frac{(1 - F_{wo}(x, \mu, \xi, \sigma))^2}{2} \in \langle 0;0,5 \rangle . \quad (9)$$

In Eq. (9) P_{jitter} reaches the maximum of 0,5 which shows an extreme, that every packet is mixed. From every pair one packet would be discarded and additional packet loss on jitter buffer reaches theoretical maximum of 50 % (transmittance factor 50 % equal to behavior of binary symmetric channel).

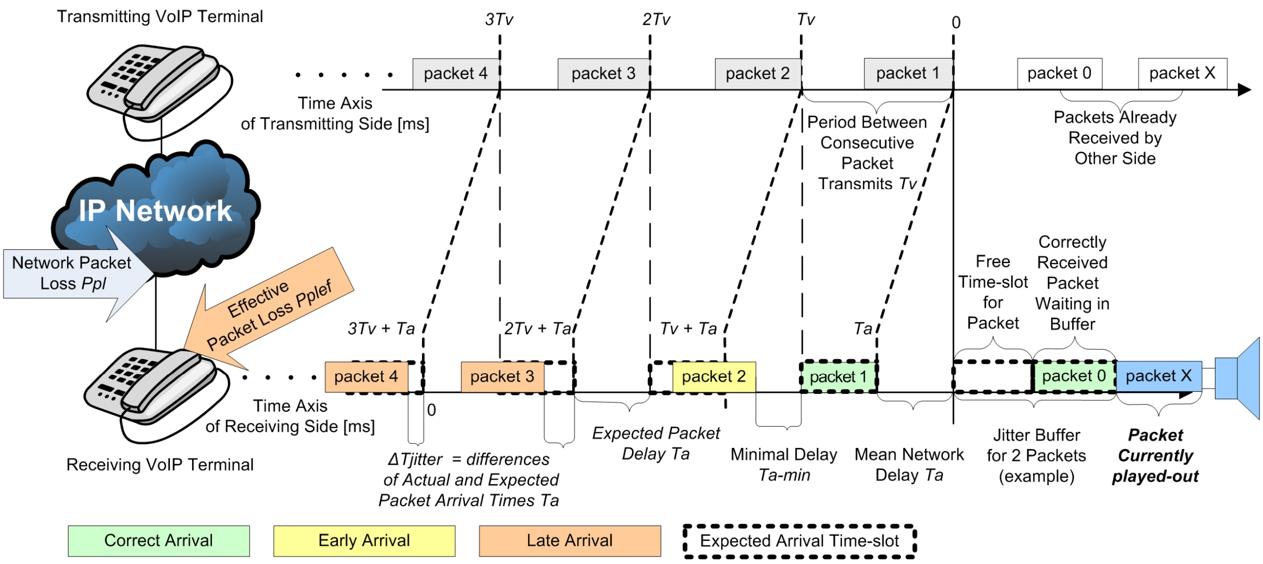


Fig. 4: Packet transmission delay, arrival time and jitter buffer function.

4. Simulation and Measurement of Jitter Effects

Measurements were carried out to model and prove jitter buffer behaviour and to find correct shape parameter ξ for Eq. (3) and (4) to find best fitting function modelling jitter buffer packet loss *Pjitter*. We have simulated MOS dependence of following codecs on QoS parameters:

- G.711 with PLC and without PLC, 64 kb/s, 40 ms jitter buffer, 20 ms/packet,
- G.723.1 ACELP, 5,3 kb/s, 60 ms jitter buffer, 30 ms/packet,
- G.723.1 MPMLQ, 6,3 kb/s, 60 ms jitter buffer, 30 ms/packet,
- G.726, 32 kb/s, 40 ms jitter buffer, 20 ms/packet,
- G.729, 8 kb/s, 40 ms jitter buffer, 20 ms/packet.

Key QoS network parameters were chosen as follows:

- One-way end-to-end delay $Ta \in \{0, 20, 50, 100, 150, 200, 300, 400\}$ [ms],
- Network packet loss $Ppl \in \{0, 1, 2, 3, 5, 7, 10, 15, 20\}$ [%],
- Network jitter of 20, 40 and 80 ms with Pareto distribution.

Special set of measurement and simulation was dedicated to statistics testing and was performed with WANem, EasyFit and StatAssist software suite as follows:

- Generic codec without PLC (valid for PCM or ADPCM codecs family G.711 and G.726) with jitter buffer from 2 to 6 packets (when packetization = 20 ms, then the buffer would be from 40 to 120 ms) and RFC 1889 jitter from 0 to 100 ms in steps of 10 ms. Delay was fixed to 100 ms and network packet loss P_{pl} set to '0' to show pure influence of buffer on resultant real loss.

VoIP traffic was simulated using IxChariot endpoint on dedicated computer with 100 BASE-TX Ethernet network card, switched through computer with two network cards emulating transport network by imposing QoS parameters on relayed packets by WANem software. VoIP stream was received with third computer with IxChariot endpoint.

For each codec and combination of QoS parameters values was simulated a packet flow of 10000 packets corresponding to 200 seconds of continuous VoIP call with 20 ms packetization. Altogether more than 189 combinations of parameters were simulated. Test runs, where *Ta.2 < jitter* were omitted as e.g. mean end-to-end of 20 ms with 80 ms jitter cannot exist when using

to Contents, page 1

aforementioned statistics and jitter calculation methods.

Measured dependencies of MOS on QoS parameters were processed as 3D graphs and in tabular form. Estimate of MOS for G.711 codec is through E-model, proposed modified E-model and measurement is enclosed in section 6. Data and net packet losses at jitter buffer are analyzed for particular jitter strength and buffer sizes. Resultant additional losses are in Tab. 2 and 3. Pure jitter buffer loss measurements and packet flow statistic tests are present in section 7.

Tab.2: Measured jitter buffer loss without reordering capability.

Jitter [ms]	**Jitter buffer size [ms]**	**Average loss on jitter buffer P_{wo}(jitter,buffer-size)**	**Complement to loss – transmittance of buffer F_{wo}(jitter,buffer-size)**
20	40	0,116114	0,883886
20	60	0,029393	0,970607
40	40	0,307519	0,692482
40	60	0,180276	0,819725
80	40	0,509000	0,491000
80	60	0,399790	0,600206

Tab.3: Measured jitter buffer loss with RTP reordering capability.

Jitter [ms]	**Jitter buffer size [ms]**	**Average loss on jitter buffer P_{wr}(jitter,buffer-size)**	**Complement to loss – transmittance of buffer F_{wr}(jitter,buffer-size)**
20	40	0,001049	0,998951
20	60	0,000316	0,999684
40	40	0,059720	0,940280
40	60	0,014509	0,985491
80	40	0,147420	0,852580
80	60	0,078991	0,921009

5. Equipment Impairment Factor Calculation Modification Using Effective Packet Loss

Based on simulation results and measurements we have determined optimal shape parameter ξ giving the smallest overall MSE error of *Pwr* for all jitter and jitter buffer size combinations from Tab. 2 and 3. Optimal value for our simulation is ξ = - 0,1 with relative MSE of MOS of only 12,0 %. We add two parameters to E-model through Eq. (10) which incorporates jitter buffer size x [ms] and network jitter σ [ms].

$$P_{jitter} = \frac{\left(1-\left(1-\left(1+\frac{\xi(x-\mu)}{\sigma}\right)^{-\frac{1}{\xi}}\right)\right)^2}{2}. \qquad (10)$$

After substitution of parameters ξ = -0,1and μ = 0 we get equation for jitter buffer packet loss as Eq. (11):

$$P_{jitter} = \frac{\left(1+\frac{-0{,}1x}{\sigma}\right)^{20}}{2}. \qquad (11)$$

Further we rewrite Eq. (7) as Eq. (12) and substitute P_{jitter} into Eq. (12). We get Eq. (13) which expresses effective packet loss P_{plef} incorporating network and jitter buffer packet loss. P_{plef} from Eq. (13) substitutes P_{pl} in Eq. (2) which leads to Eq. (14):

$$P_{plef} = P_{pl} + P_{dejitter} + P_{pl} \cdot P_{dejitter}. \qquad (12)$$

$$P_{plef} = P_{pl} + \frac{\left(1+\frac{-0{,}1x}{\sigma}\right)^{20}}{2} - P_{pl} \cdot \frac{\left(1+\frac{-0{,}1x}{\sigma}\right)^{20}}{2}. \qquad (13)$$

$$I_{e,eff} = I_e + (95 - I_e) \cdot \frac{P_{plef}}{P_{plef} - B_{pl}}. \qquad (14)$$

Equation (14) is the final proposed equation for equipment impairment factor calculation in E-model including jitter buffer loss and jitter buffer size through effective packet loss P_{plef}.

6. Results

Proposed change in equipment impairment factor calculation leads to improved MOS estimate of E-model when network jitter is present. Network jitter effects as increased packet loss were mostly negligible when network jitter was smaller than approx. ½ of jitter buffer length. Real jitter buffer have always reordering capability through RTP packet numbering and queue without reordering serves as a basic concept for formulating appropriate equations. In our application marginal value for jitter was 20 ms, thus our simulations and measurements were conducted with 20, 40 and 80 ms jitter, see Tab. 3 for results. Discovered dependence of jitter buffer packet loss at different network jitter strengths for different buffer sizes is illustrated at Fig. 5.

to Contents, page 1

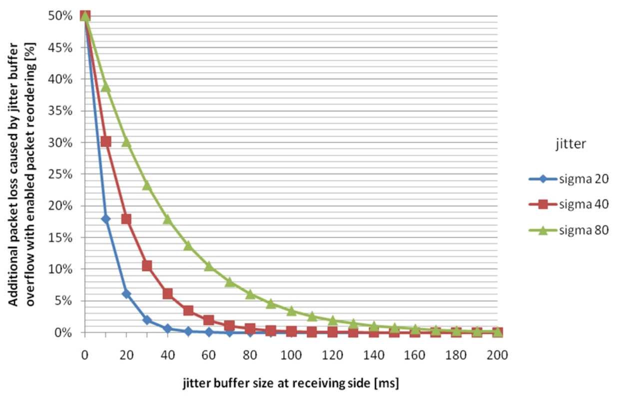


Fig. 5: Jitter buffer packet loss P_{jitter} graph for different jitter.

Table 4 summarizes MOS estimate accuracy by comparison of original and modified E-model with measurements from simulation through average differences in MOS category values. Figure 6 shows comparison of MOS estimates of original E-model, modified E-model and measurements. Offset of MOS estimate caused by presence of jitter is successfully compensated with slight local deviations caused by smaller size of statistical set of measurements. Effective packet loss and resulting end-to-end delay at receiver were greater than objectively measured network QoS parameters due to buffer behaviour and buffering delay.

Tab.4: MOS estimate improvement and comparison of estimation errors of original and modified E-model.

	Original E-model estimate		Proposed Modified E-model estimate		MOS estimate
Jitter [ms]	**MOS diff. (absolute)**	**MOS diff. (%)**	**MOS diff. (absolute)**	**MOS diff. (%)**	**Modified model has**
G.711 codec without PLC, 64 kb/s					
20	-0,17	-8,8	-0,01	-2,5	**improved**
40	-1,03	-37,6	-0,14	-9,9	**improved**
80	-1,12	-39,8	0,01	-2,9	**improved**
G.711 codec with PLC, 64 kb/s					
20	0,00	0,5	0,06	2,2	**worsened**
40	-0,64	-19,4	-0,16	-6,1	**improved**
80	-1,43	-45,3	-0,32	-17,2	**improved**
G.723.1 ACELP, 5,3 kb/s					
20	-0,37	-14,6	-0,35	-13,8	**improved**
40	-0,52	-20,3	-0,33	-13,9	**improved**
80	-1,06	-41,7	-0,33	-20,1	**improved**
G.723.1 MPMLQ, 6,3 kb/s					
20	-0,25	-9,3	-0,23	-8,5	**improved**
40	-0,41	-15,6	-0,19	-8,0	**improved**
80	-1,00	-38,8	-0,16	-12,2	**improved**
G.726 codec, 32 kb/s					
20	-0,89	-32,4	-0,83	-30,7	**improved**
40	-1,43	-49,2	-0,88	-37,4	**improved**
80	-1,75	-58,5	-0,93	-43,7	**improved**
G.729 codec, 8 kb/s					
20	0,04	3,1	0,12	6,2	**worsened**
40	-0,63	-21,2	0,02	2,1	**improved**
80	-1,45	-53,0	-0,52	-32,5	**improved**

Proposed ITU-T E-model calculation modification improved MOS estimate significantly in most cases. Further simulations and improvements to emulation setup would obtain statistically more reliable base data for function fitting with more data points.

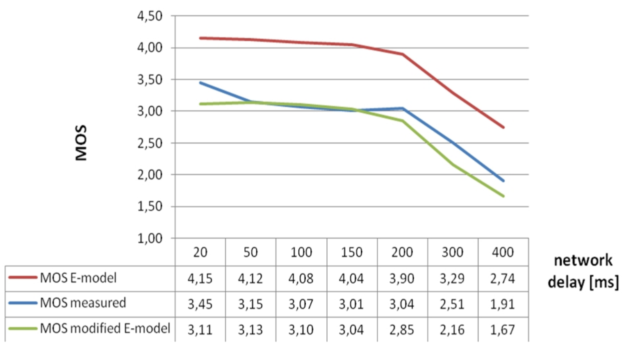


Fig. 6: Comparison of MOS estimate for G.729 codec at 40 ms jitter, 40 ms jitter buffer, 0 % packet loss at various network delay.

Potential improvements in MOS estimation by adding extra parameter describing burstiness of packet stream are further tested in next section by imposing 0 % delay correlation (absolutely random Pareto-distributed packet stream) and 50 % delay correlation to obtain maximum burstiness across larger time-scale still maintaining Pareto-distributed delay.

7. Experiment, Statistic Tests

In this experimental part of the paper we present measurements and statistical tests of VoIP input streams generated according to the most generally accepted knowledge of general IP packet streams [8], [9], [10], [12] to support our model proposed in previous sections.

7.1. Relation between Jitter, Buffer Size and Additional Packet Loss on Jitter Buffer

Figure 7 shows dependency of packet loss occurring on jitter buffer under jitter with varying intensity. Buffer size is normalised relative to jitter amount. It means that if jitter buffer is 2 packets and jitter intensity is 1,5 packets, for 20 ms packetization it would refer to 40 ms jitter buffer and 30 ms RFC 1889 network jitter and vice versa.

to Contents, page 1

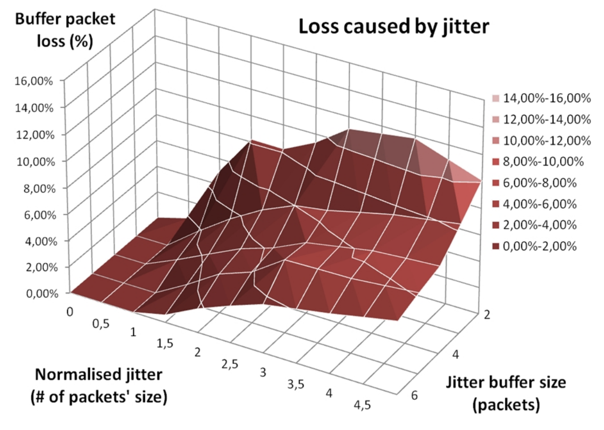


Fig. 7: Packet loss on buffer on relative buffer size to jitter.

7.2. MOS, Jitter and Buffer Size Relations

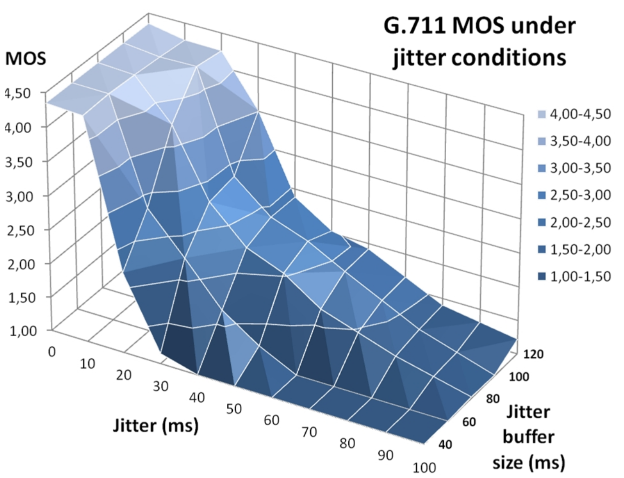


Fig. 8: Actual MOS performance of G.711 codec without PLC under varying jitter and buffer size as measured on Pareto stream.

7.3. Test of Packet Interarrival Distribution

We have chosen Pareto distribution for its long-tailed property for modelling packet input streams from intranet / internet network based on several rigorous studies of IP traffic [8], [9], [10], [12]. Packet interarrival times are defined as timestamp differences. To keep statistical moments as median and variance finite [12], we abstract from network by defining two losses – on the network and on the buffer - as described in Eq. (7) and (8). The input packet stream in size of 10000 packets with Pareto-distributed delay was generated by IxChariot. Stream was sent through WANem, which applied Pareto delay distribution and various jitter to incoming packet stream. The stream was captured by Wireshark packet sniffer at receiver side along with IxChariot performing MOS estimate. RTP timestamp data were loaded into spreadsheet. Then the interarrival times, the first differences of packet RTP timestamps, were calculated according to RFC 1889 recommendation (Eq. 6). Iterative distribution fitting was performed using various distributions to find best fit parameters. These parameters and distributions were put under Kolmogorov-Smirnov, Anderson Darling and Chi-Squared tests to find best descriptive statistics of Pareto-distributed stream time differences with applied jitter. Results of finding best descriptive statistics with optimal iteratively found parameter set with error of 10e-5 are sorted in Tab. 5.

Table 6 summarizes best fitting distributions based on aforementioned three statistical tests on dataset. The quest was to test, whether Pareto distribution was suitable not only to describe packet delay distribution of each individual VoIP stream extracted from larger traffic, but also its differences – interarrival time distribution affected by equally distributed jitter and various delay correlation (i.e. burstiness) averaged over timescale of thousands of packets – the magnitude corresponding to common call duration times.

Tab.5: Best fit parameters of tested distributions.

Distribution	Best fit distribution parameters
Generalized Pareto (GPD)	*k=0,19328; σ=0,0224; μ=-0,00306*
Generalized Extreme	k=0,36239; σ=0,01384; μ=0,00909
Weibull	α=0,39981; β=0,01718
Gen. Gamma	k=0,98444; α=0,40502; β=0,05293
Gamma (3P)	α=0,3377; β=0,08127; γ=1,3000e-5
Log-Pearson 3	α=6,081; β=-1,175; γ=1,641
Generalized Gamma (4P)	k=1,6025; α=0,20097; β=0,09338; γ=1,3000e-5
Laplace	λ=39,109; μ=0,0247
Weibull (3P)	α=0,4745; β=0,01408; γ=1,3003e-5
Gamma	α=0,46676; β=0,05293
Logistic	σ=0,01994; μ=0,0247
Lognormal	σ=2,8971; μ=-5,504

Tab.6: Best fit distribution ranking.

Distribution	Kolmogorov-Smirnov Test		Anderson Darling Test		Chi-Squared Test	
	Stat.	Best Fit Rank	Stat.	Best Fit Rank	Stat.	Best Fit Rank
Gen. Pareto	*0,12738*	*1*	*71,243*	*1*	*574,28*	*2*
Gen. Extreme	0,13416	2	86,466	2	377,48	1
Weibull	0,15924	3	188,58	4	2293,7	8
Gen. Gamma	0,16413	4	217,81	6	N/A	N/A
Gamma (3P)	0,16819	5	237,21	8	N/A	N/A
Log-Pearson 3	0,17897	6	220,33	7	2531,8	12
Gen. Gamma (4P)	0,1814	7	186,34	3	1798,2	7

to Contents, page 1

Laplace	0,19224	8	274,75	9	574,51	3
Weibull (3P)	0,19363	9	345,11	13	N/A	N/A
Gamma	0,20779	10	207,83	5	1675,6	6
Logistic	0,22469	11	280,74	10	668,76	4
Lognormal	0,22694	12	343,23	12	4324,8	13

Statistical tests showed as a proof of concept, that GPD Pareto distribution is also the most suitable one for describing interarrival times of general long-tailed LAN/WAN packet streams impaired by random jitter with equal distribution. This shows also Pareto distribution to be the best compromise between calculation complexity (compared to fractal modelling methods) and statistical significance for modelling also jitter buffer loss behaviour under variable jitter conditions. Dataset CDF is illustrated in Fig. 9 and represents normalised jitter buffer packet transmittance under chosen jitter (in our case 40 ms) dependent on jitter buffer size on x-axis (from 0 to 1 second = 0 to 1000 ms).

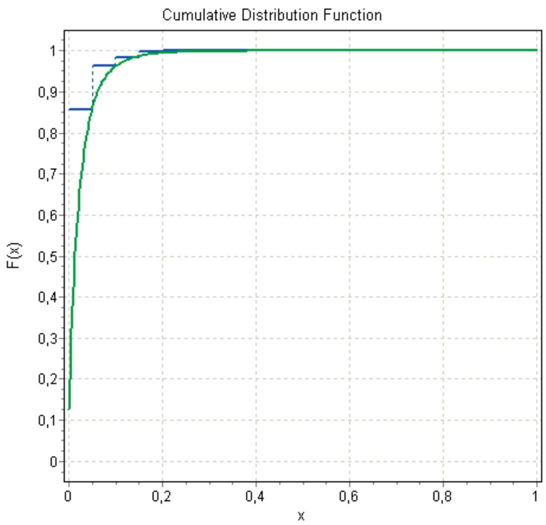


Fig. 9: GPD distribution fitting to real network measured interarrival times on CDF plot: green line – GPD CDF (best fit, parameters from Tab. 5); blue histogram – real interarrival times.

To explain best fit parameters of GPD from Tab. 5, the relation between these parameters to parameters in Eq. (9), (10), (11) and (13) is following:

- σ in Eq. (9) = std. deviation of bounded GPD Pareto realization corresponds to optimised σ in Tab. 5. Proposed relation between σ and actual jitter J substituted can be expressed in ratio $J/\sigma \in$ <1;2>. For actual imposed 40 ms network jitter the optimized parameter was σ=0,0224 (s) = 22,4 ms what would yield J/σ ratio = 22/14 $\in$ <1;2>. Actual parameter substitution ratio needs further testing and finding acceptable compromise for various jitter sizes,
- ξ in Eq. (9) = shape parameter and corresponds to optimised $-k$ in Tab. 5. Actual shape parameter for our preliminary model in Eq. (9), (10), (11) and (13) was chosen to be $\xi = \lfloor -k \rfloor$ rounded to one tenth in order to maintain exponent in Eq. (9), (10), (11) and (13) an integer. The resulting error in jitter buffer loss estimate does not exceed 1 % in our case, but further optimisation and testing could lead to better approximation while maintaining simplicity of calculation,
- μ in Eq. (9) = location parameter corresponds to μ= – 0,00306 in Tab. 5. Actual shape parameter for our preliminary model in Eq. (9), (10), (11) and (13) was chosen to be $\mu = 0$ with negligible effect.

Selected common distributions with their best fit parameters were chosen to show main reason of their unsuitability in packet interarrival time and jitter buffer behaviour estimates as visualised in Fig. 10 below.

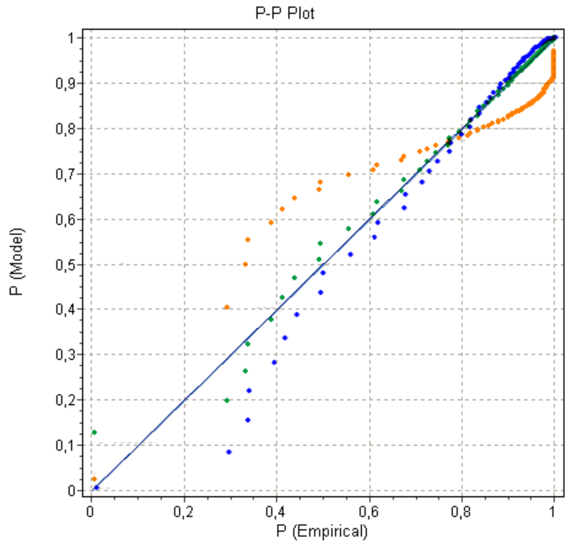


Fig. 10: Comparison of fitting on interarrival time on P-P plot for several distributions: orange – Log-Normal; blue – Exponential (2P); green – Generalized Pareto (best fit).

7.4. Burstiness Effect on Jitter Buffer Loss

Analysed set of data showed no significant effect of interarrival time correlation, i.e. burstiness on proposed modified E-model estimate. The most significant property is the statistical distribution of input stream that if maintained from long-term view (order of hundreds of packets) leads to stable per call MOS estimate. Measured difference with and without correlation was less than 0,5 % average for all combinations of jitter and buffer sizes tested. Difference can be seen in Fig. 11 as a crossection of red (buffer loss on RTP stream with delay correlation of 50 %) and blue plane (correlation = 0 %).

to Contents, page 1

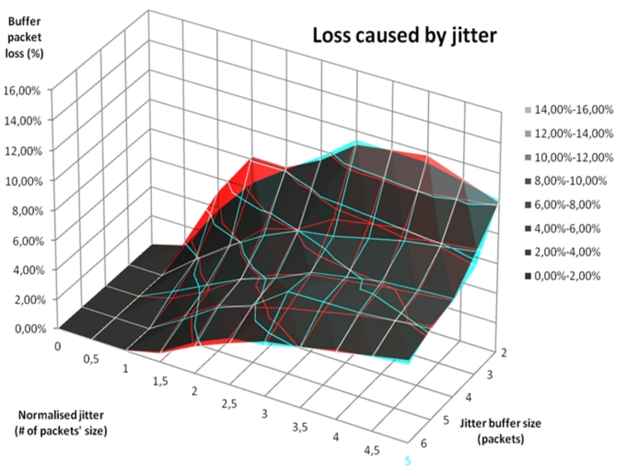


Fig. 11: Difference of jitter buffer packet loss on 50 % bursty (red) and random uncorrelated (blue) Pareto packet stream.

8. Conclusion

Proposed E-model parameter modification by substitution of actual network packet loss yield statistically improved and more coherent MOS VoIP call quality estimates extending to larger network packet losses and greater overall packet losses caused by real network with jitter. Jitter induced packet loss cannot be supported by RTP and RTCP protocols as one of the QoS parameters, because it is purely a matter of user end device, not the network part. Our testing showed great significance and importance of accounting for receiving jitter buffer behavior, for it affect perceived VoIP call quality more than network itself. Proposed modification can be applied generally on WAN or LAN networks and can describe RTP VoIP input stream from various networks with high enough statistical significance to model buffer behavior under Pareto/D/1/K input stream and improve E-model estimate over wide range of jitter and buffer sizes.

Acknowledgements

This paper and whole underlying research has been sponsored by national research projects VEGA 1/0565/09 and SMART II: ITMS 26240120029.

References

[1] ITU-T G.107. The E-model, a computational model for use in transmission planning, *ITU-T Recommendation G.107, ITU-T.* Geneva, Switzerland: May 2000.

[2] ITU-T G.107. The E-model, a computational model for use in transmission planning, *ITU-T Recommendation G.107, ITU-T.* Geneva, Switzerland: 2002.

[3] ITU-T G.107 – Amendment 1, Provisional Impairment Factor Framework for Wideband Speech Transmission, *ITU-T Recommendation G.107, ITU-T.* Geneva, Switzerland: June 2006.

[4] ITU-T G.109. Definition of categories of speech transmission quality, *ITU-T Recommendation G.109, ITU-T.* Geneva, Switzerland: September 1999.

[5] ITU-T G.133. Appendix I: Provisional planning values for the equipment impairment factor Ie and packet-loss robustness factor Bpl, *ITU-T Recommendation G.133, ITU-T.* Geneva, Switzerland: September 1999.

[6] ITU-T STUDY GROUP 12 – DELAYED CONTRIBUTION 106: *Estimates of Ie and Bpl parameters for a range of CODEC types.* Geneva, Switzerland: 27-31 January 2003.

[7] KOH, Y.; KISEON, K. Loss probability behavior of Pareto/M/1/K queue. *Communications Letters, IEEE*, January 2003. Vol.7, No.1, p. 39-41. ISSN 1089-7798.

[8] KASAHARA, S. Internet traffic modeling: a Markovian approach to self-similar traffic and prediction of loss probability for finite queues. IEICE *Transactions on Communications: Special Issue on Internet Technology*. 2001. Vol. E84-B, No. 8. pp. 2134-2141. ISSN 1745-1345.

[9] MIRTCHEV, S.; GOLEVA, R. Evaluation of Pareto/D/1/K Queue by Simulation. In *International Book Series "Information Science and Computing"*. Technical University of Sofia. Sofia, Bulgaria: June 2008. pp. 45-52. ISSN 1313-0455.

[10] DANG, T., D.; SONKOLY, B.; MOLNÁR, S. Fractal analysis and modeling of VoIP traffic. In *Telecommunications Network Strategy and Planning Symposium. NETWORKS 2004, 11th International conference*, 13-16 June 2004. pp. 123- 130. ISBN 3-8007-2840-0.

[11] KLÚČIK, S.; TISOVSKÝ, A. Queuing Systems in Multimedia Networks. *Elektrorevue*. Vol. 15, 16.11.2010 (2010), No 99. ISSN 1213-1539.

[12] CROVELLA, M. E.; BESTAVROS, A. Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes. *IEEE/ACM Transactions on Networking*, vol. 5, No. 6, December 1997. ISSN 1063-6692.

[13] TIJMS, H., C. Approximations for the Overflow Probability in Finite Buffer Queues, *Serie Research Memoranda*. Amsterdam: Faculteit der Ekonomische Wetenschappen en Econometrie, August 1992.

[14] ZHANG, W.; HE, J. Statistical Modeling and Correlation Analysis of End-to-End Delay in Wide Area Networks. In *Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*. IEEE, 2007. Vol. 3, pp. 968 – 973. ISBN 978-0-7695-2909-7.

[15] JIANG, W.; SCHULZRINNE, H. Comparison and Optimization of Packet Loss Repair Methods on VoIP Perceived Quality under Bursty Loss. In *NOSSDAV '02 Proceedings of the 12th international workshop on Network and operating systems support for digital audio and video*, May 12-14, 2002, Miami Beach, Florida, USA. pp. 73-81. ISBN 1-58113-512-2.

About Authors

Adrian KOVAC, born in 1986, has received master's degree in Telecommunications in 2010 at the Slovak University of Technology, Faculty of Electrical

to Contents, page 1

Engineering and Information Technology in Bratislava, where he currently pursues his Ph.D. studies and provides lectures about communication network security and services. In 2008 he received Bachelor's degree in Business Informatics at the University of Economy in Bratislava. Adrian specialises in network security, performance testing and modelling of communication systems behaviour. His current research topic is modelling and measurement of VoIP call quality and secure communication.

Michal HALAS was born in 1978 in Slovakia. He graduated from Slovak University of Technology and received the electronic engineering degree in 2003. Since this year he has started postgraduate study at Department of Telecommunications STU Bratislava. In 2006 he received his Ph.D. from Slovak University of Technology Nowadays he works as a lecturer in Department of Telecommunications of FEI STU in Bratislava and topics of his research interests are Speech quality, Next Generation Networks and IP telephony.

Milos ORGON was born in 1956. He received his Master degree and Ph.D. degree from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava in 1980 and 1988, respectively. Nowadays he works as an associate professor at the Department of Telecommunications of FEI STU Bratislava. He has been engaged in the research and development of telecommunication networks and services in liberalized environment for the area of convergent technologies.

Miroslav VOZNAK was born in 1971. He has been studying Telecommunications engineering at VSB-TU Ostrava and he holds a doctorate in Telecommunications. He received Ph.D. degree in 2002 at the Faculty of Electrical Engineering and Computer Science where he works as an associate professor nowadays. He gives the lectures at the Department of Telecommunications of VSB – TU in Ostrava.

to Contents, page 1

COMPARISON OF CURRENT FRAME-BASED PHONEME CLASSIFIERS

Vaclav PFEIFER[1], *Miroslav BALIK*[1]

[1]Department of Telecommunications, Faculty of Electrical Engineering and Communication, Brno University of Technology, Purkynova 118, 612 00 Brno, Czech Republic

pfeifer@feec.vutbr.cz, balik@feec.vutbr.cz

Abstract. *This paper discusses current approaches for frame-based classification and evaluates today's most common frame-based classifiers. These classifiers can be divided into the two main groups – generic classifiers which create the most probable model based on the training data (for example GMM) and discriminative classifiers which focus on creating decision hyper plane (SVM based methods). A lot of research has been done with the generic classifiers and therefore this paper will be mainly focused on the discriminative classifiers. Four discriminative classifiers are presented – two linear and two non-linear. All of these discriminative classifiers implement a hierarchical tree root structure over the input phoneme group which shown to be an effective. Moreover, two efficient training algorithms are presented. First, we demonstrate advantages of discriminative classifiers by comparison with a standard generic classifier represented by a GMM. Second, we show benefits of our proposed training algorithm. All tests were performed for English only - over the TIMIT speech corpus (corpus of Native American speakers).*

Keywords

Classifier, comparison, frame-based, phoneme, speech, hierarchical.

1. Introduction

Phoneme classification is a task of deciding the identity of an unknown speech utterance (mostly short ones) [1]. The correct classification plays important role in most of the current state-of-the-art speech systems – for example speech recognition, spoken term detection, etc. Based on the classifier input, the classifiers can be further divided into sequence based classifiers and frame based classifiers. Most of the current speech processing systems are based on the sequence based classifiers. The sequence modeling is performed using a Hidden Markov modeling (HMM) [2], [3] and classification itself can be done using Gaussian mixture modeling (GMM) [4], neural network (NN) [5] or support vector machines (SVM) [6]. These systems are mostly denoted as a HMM/GMM or HMM/NN, etc. [4], [7], [8]. The main advantages of these systems are simple phoneme modeling and good output results (especially for the phoneme posteriors) [9]. The disadvantage of the HMM-based systems lies in the Baum-Welch (BW) training algorithm which is known for his convergence to local maxima. This problem is mostly solved by multiple algorithm initializations which can be extremely time-consuming. Another problem is that these algorithms do not aim on minimization some objective function (e.g. specific loss function) [10].

Few researchers proposed different classifiers with the different results. For example authors in [11] devise a naive Bayes classifier based on the reconstructed phase space. Authors in [5] or [12] proposed new type of feature extraction technique. Most of these approaches do not aim on the improving acoustic models (AM) but rather on defining more sophisticated feature extraction techniques etc. [5], [12]. In the recent years large margin and kernel methods have proven to be an effective tool for the tasks of speech processing (e.g. speech recognition, keyword detection etc.). Most of these systems aim on the acoustic models improving with the use of proper frame-based phoneme classifiers [13], [14]. The frame-based classification is a task of deciding the identity of the each speech frame (typically 25 ms length). Due to the lack of sequence modeling these systems are less accurate compared to the sequence based. The advantage is in the proper application with the specific system, like [10], [14].

Based on the recent advantages in large margin and kernel methods and pioneering research of O. Dekel and J. Keshet [6], [10] this paper presents a simple frame-based linear phoneme classifier. The main idea is in the definition of so called prototype functions for each of the phoneme and the decision is made according to their

to Contents, page 1

similarity to each of these functions. Furthermore we have incorporated a distance metric for the hierarchical structure like in Fig. 1. This metric represents a tree induced error over the hierarchical structure and costs

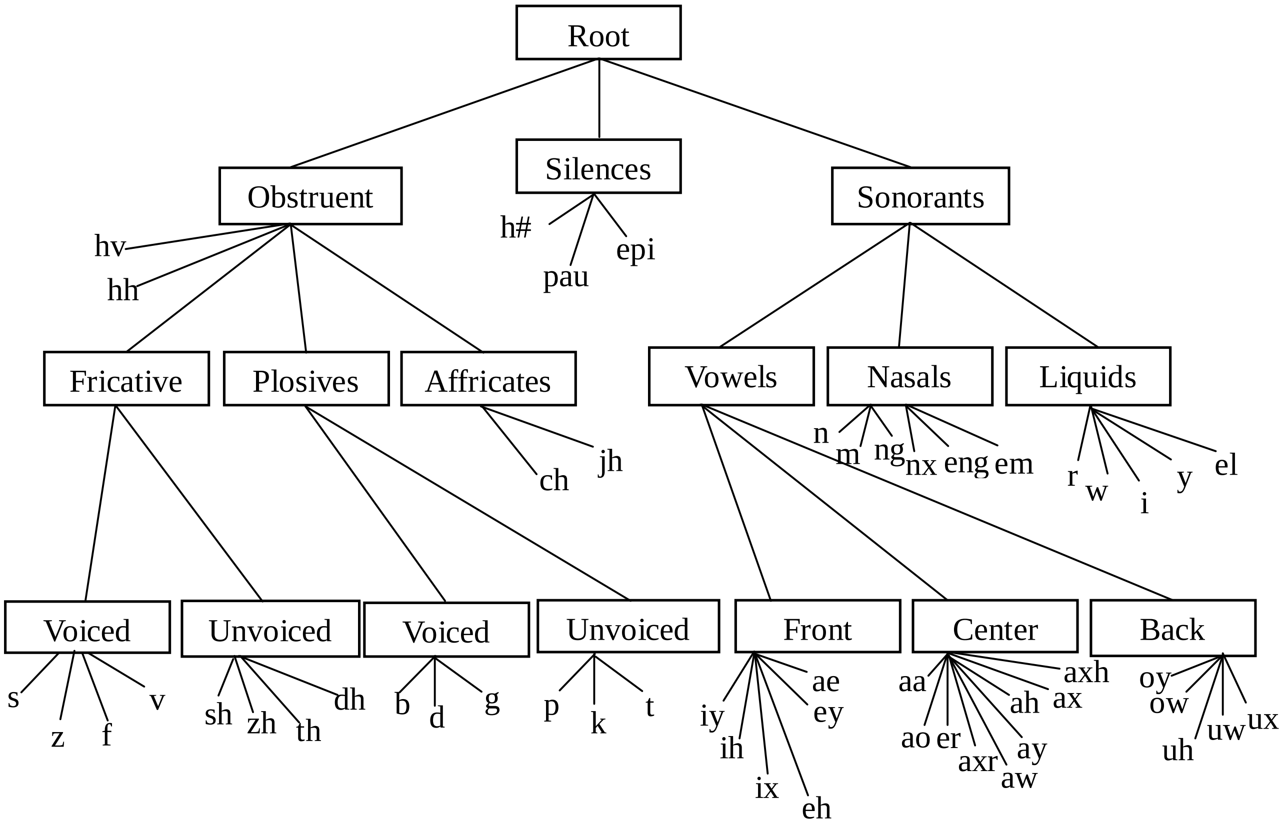


Fig. 1: TIMIT phonetic TREE structure.

misclassified phonemes according to their distance in the tree [6]. For example, classify an utterance as a phoneme *iy* instead *ih* is less severe than phoneme *uw* as a *w*. The hierarchical structure in Fig. 1 represents phonetic tree structure for the American English and a set of the used phonemes is derived from the TIMIT speech corpus [15]. Given the fact that current speech corpora contains large amounts of annotated speech samples we have proposed an efficient learning procedure for the prototype function estimation. This learning algorithm is based on the batch generalization and the results clearly show the benefit of proposed algorithm. Moreover we have proposed a batch-classifier for the whole phoneme sequence classification. The advantage of our approach is that the efficient learning algorithm can be used both for the frame-based classification and whole sequence-based classification.

For the task of evaluation two different metrics are used. First metric defines a phoneme error rate (PER) which corresponds to the number of misclassified phonemes. Second metric defines number of misclassified phoneme groups (MISS). For example to classify an utterance as *iy* instead of *ih* is a mistake for the PER metric while the metric MISS evaluate correct classification.

This paper is organized as follows. In section 2 we define the problem settings. In section 3 we present a classification rule. Section 4 and 5 introduce our proposed algorithms. In section 6 an evaluation is performed and section 7 concludes our results.

2. Problem Settings

Let $\mathbf{x}$ be the sequence of acoustic feature vector, so $\mathbf{x} = (x_1, x_2 \ldots x_T)$, $x_T \in \mathbf{X}$, where $\mathbf{X} \subset \mathbf{R}n$ is the acoustic feature domain. Let $\boldsymbol{Y}$ be a set of phonemes and phoneme groups defined according the hierarchical structure like in Fig. 1. Let us further consider align between each of the phoneme or phoneme group $\boldsymbol{y} \in \boldsymbol{Y}$ and appropriate acoustic features $\mathbf{x} \in \boldsymbol{X}$. Denote T to be a corresponding hierarchical structure (like the one in Fig. 1). The number of all vertices in the tree structure T is denoted as a $k = |\boldsymbol{Y}|$, in other words k encompasses the number of all phoneme and phoneme groups so $\boldsymbol{Y} = \{0,\ldots, k-1\}$, where 0 represents a tree root of the hierarchical structure $\boldsymbol{T}$ [6].

Let us define a metric $\gamma(\cdot, \cdot)$ over this hierarchical tree structure $\boldsymbol{T}$ as a number of all edges (unique path) between two different phonemes of phoneme groups u, v. For any pair of phonemes u, v let $\gamma(u, v)$ be their distance

to Contents, page 1

in the tree root structure T, while following triangle equality holds $\gamma(u, v) = \gamma(v, u)$ and $\gamma(u, u) = 0$ since $\gamma(u, v)$ is a non-negative function. based on the stated definitions let us further define a tree induced error $\gamma(u, v)$ as a unique path from the phoneme u to v, so the tree induced error incurs only while predicting incorrect phoneme [10].

For every phoneme and phoneme group (expect the tree root) $v \in \{Y\backslash 0\}$ we denote $A(v)$ to be a parent of v in the T. Further we define an ancestor of v as a $A^{(i)}(v)$ which is recursively defined as follows,

$$A^{i}(v) = A\left(A^{(i-1)}(v)\right), \tag{1}$$

and $A^{(0)}(v) = v$. For each phoneme and phoneme group $v \in Y$ we define $P(v)$ to be a number of vertices from phoneme v to the tree root 0 resp. $P(v)$ encodes a unique path from phoneme v to tree root 0 [6], [10],

$$P(v) = \left\{u \in Y : \exists i, u = A^{(i)}(v)\right\}. \tag{2}$$

The goal of the frame–based classifier is to determine frame identity – to decide which the most probable phoneme that frame belongs to. In case of the hierarchical based classifier there could be stated another goal. To determine frame's phoneme group identity. Both of these stated goals can be measured by above mentioned metrics (PER and MISS).

3. Classification Rule

The proposed frame-based classifier (resp. classification function) f: $X \rightarrow Y$ makes its prediction according to the input set of prototypes (weights vectors) W defined for each of the phoneme and phoneme group v. Each of the prototype W can be any vector in Rn and our goal is to train frame-based classifier f which attains low tree induced error $\gamma(u, v)$ on the training samples. For the frame-based training algorithm the input training database is defined in the following form $S = \{(xi, yi)\}m$ i=1, where m is the number of all training samples (pairs), that is set S consists of m pairs in the following form $x_i \in X$ and $yi \in Y$ so the training is performed per each of the frame (therefore frame-based). The task of learning is then simplified to find appropriate weights vectors $W_1 \ldots Wk$-1. Linear frame-based classifier f is defined according the following formula:

$$f(x) = \underset{v \in Y}{\arg\max}\left(W^{v} \cdot x\right). \tag{3}$$

The classifier defined by the Eq. (3) does not include hierarchical structure T. To incorporate tree root structure T into the classification function we have to define a new set of weights vectors w,

$$w^{w} = W^{v} - W^{A^{1}(v)}, \tag{4}$$

so we had decided to work with the partial differences w rather that with standard weights W. The weight vector W can be furthermore rewritten based on the Eq. (4) as follows,

$$W^{v} = \sum_{u \in P(v)} w^{v}. \tag{5}$$

Based on the Eq. (5) and (3) the resulting hierarchical frame-based classifier can rewritten into the form of Eq. (6),

$$f(x) = \underset{v \in Y}{\arg\max} \sum_{u \in P(v)} w^{u} \cdot x. \tag{6}$$

4. Efficient Training Algorithm

The proposed training algorithm is based on the concept of O. Dekel [6] and J. Keshet [10] and both of these algorithms are based on the frame-based classification function (like the one in Eq. (3)) so the learning procedure is also proposed as a frame-based. The principle of our learning algorithm is based on the idea of sequential training and sequence generalization. On each round not a simple frame updates our weights vectors w but rather the whole generalized phoneme frames sequence. In other worlds, on each round all the appropriate weights w of each phoneme v are updated at once. Furthermore, our derived prototypes w_v can still be efficient in a frame-based classification as well in whole sequence (batch) classification. The principle of our learning algorithm lies in the redefinition of the classification function f defined by the Eq. (6). As stated above, this function is defined to be a frame-based so to incorporate a whole sequence prediction we have to rewrite our classification function as follows,

$$f(x) = \underset{v \in Y}{\arg\max} \sum_{u \in P(v)} mean(w_i^{u} \cdot x), \tag{7}$$

where operator mean() represents an average of the partial values xj, where j is a parameter index (e.g. first MFCC coefficient) and w_i is a weight vector in the i-th iteration step. In the theory of the Large margin and kernel methods we assume that there exists a set of prototypes $\{w(v)\}v_{\mathrm{in}Y}$ such that for each pair (x_i, y_i) and every $r \neq y_i$ the following inequality holds:

$$\sum_{v \in P(y_i)} \left\|w_i^{v} \cdot x\right\| - \sum_{u \in P(r)} \left\|w_i^{u} \cdot x\right\| \geq \sqrt{\gamma(y_i, r)}, \tag{8}$$

where y_i is a correct prediction according the classification function defined by Eq. (7) and $\|\cdot\|$ refers to L2 norm. According to the Eq. (8) we require that the difference between the correct prediction and any incorrect prediction is at least square-root of the tree-based distance between them [10]. The goal of the proposed algorithm is to find a set of prototypes which

to Contents, page 1

fulfills the margin requirement defined by Eq. (8) while incurring minimal tree-induced error [6]. In machine learning we do not minimize Eq. (8) directly but rather employs a convex hinge-loss function $\ell(\{w_i(v)\}, x_i, y_i)$

$$\ell = \left[\sum_{v \in P(y)} \left\| w_i^v \cdot x \right\| - \sum_{u \in P(y_i)} \left\| w_i^u \cdot x \right\| + \sqrt{\gamma(y_i, y)} \right]_+ , \quad (9)$$

where $[z]+ = \max\{z, 0\}$. Let us assume that there was a prediction mistake b y_i on round i and we would like to modify a set of prototypes $\{w_i(v)\}$ so the constraints defined by Eq. (8) holds. However a simple analytical solution does not exist so we introduce a simple optimization problem frequently used in SVM and machine learning theory [16]. Formally, the new set of prototypes $\{w_{(i+1)}(v)\}$ is the solution of the following optimization problem,

$$\begin{gathered} \min_{\{w^v\}} \frac{1}{2} \sum_{v \in Y} \left\| w^v - w_i^v \right\|^2 \\ s.t. \sum_{v \in P(y_i)} \left\| w_i^v \cdot x \right\| - \sum_{u \in P(r)} \left\| w_i^u \cdot x \right\| \geq \sqrt{\gamma(y_i, r)} \end{gathered} . \quad (10)$$

Note, that only the weights $\{w_i(v)\}$ defined by the path $P(v)$ are updated at each iteration. The Fig. 2 demonstrates this update – only the vertices depicted by the solid lines are updated at once.

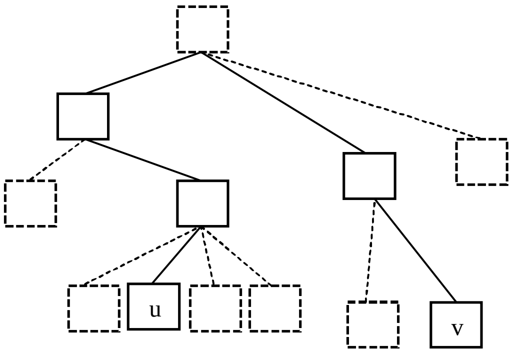


Fig. 2: Re-estimation of the weights vectors – only the solid lines are updated.

The solution to the optimization problem defined by the Eq. (10) is based on the dual representation in the form of Lagrangian [16]. We set the derivate of Lagrangian $\{w(v)\}$ to zero so the new weights $\{w_{(i+1)}(v)\}$ are estimated according the following formulas,

$$w_{i+1}^v = w_i^v + \alpha_i mean(x), \quad v \in P(y_i) \setminus P(\hat{y}_i), \quad (11)$$

$$w_{i+1}^v = w_i^v - \alpha_i mean(x), \quad v \in P(\hat{y}_i) \setminus P(y_i), \quad (12)$$

where the Lagrangian multipliers α_i are simply computed as a

$$\alpha_i = \frac{\ell(\{w_i^v\}, x_i, y_i)}{\gamma(y_i, \hat{y}_i) \cdot \|x\|_N}, \quad (13)$$

where $\|\cdot\|_N$ represents a matrix form defined by the following term

$$\sum_{i=1}^{\max(i)} \sqrt{\sum_{j=1}^{\max(j)} x_{i,j}^2} . \quad (14)$$

On each round we have generated a new set of prototypes $\{wv\ i\}$ so while the training corpus S contains m training samples we have m sets of prototypes. The last set of weight vectors $\{w_m(v)\}$ should be the best resulting prototypes but in practice an averaged weights vectors shows to be more efficient. The resulting prototypes are defined as follows,

$$w_{avg}^v = \frac{1}{m+1} \sum_{i=1}^{m+1} w_i^v . \quad (15)$$

INITIALISATION: $\forall v \in Y$: wv 1 = 0

For i=1, 2. . . m

- Algorithm receive acoustic features vector _x*i* for the phoneme *yi*
- Prediction

$$f(x) = \arg\max_{v \in Y} \sum_{u \in P(v)} mean(w_i^u \cdot x)$$

- Correct phoneme yi is revealed
- In case of incorrect prediction ($\gamma(\cdot,\cdot) \neq 0$) the hinge loss function $\ell(\{w_i(v)\}, x_i, y_i)$ is computed
- Re-estimation of the weight vectors:

$$w_{i+1}^v = w_i^v + \alpha_i^v mean(x)$$

$$\alpha_i^v = \begin{cases} \alpha_i & v \in P(y_i) \setminus P(\hat{y}_i) \\ -\alpha_i & v \in P(\hat{y}_i) \setminus P(y_i) \\ 0 & otherwise \end{cases}$$

- Where

$$\alpha_i = \frac{\ell(\{w_i^v\}, x_i, y_i)}{\gamma(y_i, \hat{y}_i) \cdot \|x\|_N}$$

Fig. 3: Proposed training algorithm.

5. Non-Linear Classifiers

The proposed training algorithm can be further incorporated with the non-linear kernel transformation. The main idea lies in the vector space separation.

to Contents, page 1

Because of speech complexity not all of the frames can be linearly well separated in the input feature space R*n*. Based on the SVM theory there a non-linear transformation can be applied on the input features (both in the training and evaluation) [16]. Again, transformed features are linearly separated (non-linearly decision hyper plane can be seen in the original feature space). To define a non-linear classifier we have to rewrite our fundamental classification rule defined by the Eq. (3) in case of the linear classifier and Eq. (6) in case of the hierarchical classifier. Because this paper primary deals with the hierarchical classifiers Eq. (6) will be rewritten but the same can be applied on the linear classifiers. The resulting classifier will be in the following form:

$$f(x) = \arg\max_{v \in Y} \sum_{u \in P(v)} \sum_{i=1}^{m} \alpha_i \cdot x_i \cdot x\,, \qquad (16)$$

where α_i is i-th Lagrangian multiplier and xi is i-th training sample. This definition is valid and expressing the whole weight vector w estimation. Equation (16) can be further rewritten in the Kernel notation in the form of following Eq. (17)

$$f(x) = \arg\max_{v \in Y} \sum_{u \in P(v)} \sum_{i=1}^{m} \alpha_i \cdot K(x_i \cdot x)\,, \qquad (17)$$

where inner product between x_i and x is represented by the kernel operator K. Based on the Eq. (16) and (17) there should be clear that the whole training database (or at least Lagrangians α_i) are necessary in the classification process. Nevertheless, only non-negative Lagrangians contributes to the final result which leads to the sparse solution.

$$\ell = \left[\sum_{v \in P(\hat{y}_i)} \sum_{j<i} \left\| \alpha_j^v \cdot K(x_i, x_j) \right\| - \sum_{u \in P(y_i)} \left\| \alpha_j^v \cdot K(x_i, x_j) \right\| + \gamma(y_i, y) \right]_+ . \qquad (18)$$

5.1. Training Algorithm

The advantage of the re-definition according to Eq. (17) lies in the possibility of kernel K pre-computation. Kernel $K(x_i, x)$ can be reformulated as $K(x_i, x_j)$, where x_i and x_j are training samples where $i, j = 1 \ldots m$. Resulting matrix G (so called Gram matrix [16]) is composed with every $K(x_i, x_j)$ value and this matrix can be pre-computed just once (for the same kernel parameters).

To develop a training algorithm we have further incorporate kernel operator into the loss function ℓ. This leads to the Eq. (18), where Lagrangians $\alpha_i(v)$ are computed based on the following equation

$$\alpha_i = \frac{\ell\left(\left\{\alpha_j^v\right\}, G(j,i), y_i\right)}{\gamma(y_i, \hat{y}_i) \cdot G(i,i)}\,, \qquad (19)$$

where $\alpha_j(v)$ is j-th Lagrangian associated with the phoneme y_j and G represents Gram matrix.

We have further experimented with the mutual combination of linear and non-linear classifiers which lead to the following efficient training algorithm – see Fig. 5.

INITIALISATION: $\forall v \in Y$: $w_1(v) = 0$, $\alpha_i(v) = 0$

Pre-computation of Gram matrix **G** [optional]

$$G(i,j) = K(x_i, x_j)$$

For i=1, 2. . . m

- Algorithm receive acoustic features vector x_i for the phoneme y_i
- Prediction

$$f(x) = \arg\max_{v \in Y} \sum_{u \in P(v)} mean(w_i^u \cdot x)$$

- Correct phoneme y_i is revealed
- In case of incorrect prediction ($\gamma(\cdot, \cdot) = 0$) the hinge loss function $\ell(\{\alpha\}, G(j,i), yi)$ is computed
- Re-estimation of the weight vectors:

$$w_{i+1}^v = w_i^v + \alpha_i^v mean(x)$$

$$\alpha_i^v = \begin{cases} \alpha_i & v \in P(y_i) \setminus P(\hat{y}_i) \\ -\alpha_i & v \in P(\hat{y}_i) \setminus P(y_i) \\ 0 & otherwise \end{cases}$$

- Where

$$\alpha_i = \frac{\ell\left(\left\{\alpha_j^v\right\}, G(j,i), y_i\right)}{\gamma(y_i, \hat{y}_i) \cdot G(i,i)}$$

Fig. 4: Efficient non-linear training algorithm.

to Contents, page 1

6. Evaluation

We have performed a number of tests to evaluate our proposed training algorithm and all algorithms were evaluated over the TIMIT speech corpus [15] which is a speech corpus of annotated utterances for American English. We have divided the TIMIT sentences into the two disjoints groups – TRAIN and TEST. We have also excluded all the *SA* sentences (dialect sentences) and we have randomly generated 80 TRAIN (for the second experiment, number of training examples will vary) and 80 TEST sentences as follow – Each sentences is uttered by the different speaker, each speaker uttered one *SI* and *SX* sentence and both sets have a uniformly distributed all the dialect regions. We have also separated all training and testing sentences so performed evaluation can be considered as a speaker independent.

In the first experiment, classifiers were compared on the two different feature extraction techniques - mel-frequency cepstral coefficients (MFCC) and perceptual linear prediction coefficients (PLP) both detailed described in the literature [17]. We used 13 basic coefficients, deltas and double deltas ($\Delta+\Delta\Delta$). We have used 15 mixtures for GMM model. Furthermore, features were normalized using the CMN/CVN technique. The Tab. 1 and 2 displays the results indicating the advantage of PLP features. The first proposed training algorithm (denoted as a $\text{Hier}_{\text{sekv}}$) had been evaluated as a standard linear classifier and shown to be more accurate compared to the frame-bases training algorithm based on the [6]. Moreover, learning time was rapidly reduced. For notation, both training algorithms were evaluated based on the same frame-based classification rule defined by the Eq. (6). The second proposed training algorithm (denoted as a $\text{Hier\_kernel}_{\text{sekv}}$) incorporated a non-linear transformation represented by the kernel operator *K*. To evaluate benefits of our proposed training algorithm we had compared our nonlinear training algorithm with the non-linear training algorithm proposed in the [6] (denoted as a Hier_kernel). Finally, all hierarchical frame-based training algorithms were compared with the standard GMM frame-based classifier (like the one in [4], [7]). To assure a convergence to global optimum we have performed a number of re-estimation of the training algorithm and the one yielding the best results over the cross-validation set had been chosen for the further evaluation.

Tab.1: PER and MISS for PLP features.

Number of sentences [-]	PER [%]	MISS [%]	Training time [min]
Hier	55	31	180
$\text{Hier}_{\text{sekv}}$	53	29	**15,8**
Hier_kernel	54	29	315
$\text{Hier\_kernel}_{\text{sekv}}$	**49**	**25**	195
GMM	52	36	35

Our second experiment demonstrates a benefit of our training algorithm. Table 3 shows that with the larger number of training sentences the proposed sequence based algorithm converges to the global optimum defined by the frame-based learning algorithm. Furthermore, at the same PER and MISS results our proposed algorithm is much more time-efficient compared to the frame-based algorithm based on the [6]. Figure 3 and 4 graphically output these results.

Tab.2: PER and MISS for MFCC features.

Classifier type	PER [%]	MISS [%]	Training time [min]
Hier	56	32	182
$\text{Hier}_{\text{sekv}}$	55	31	**16,7**
Hier_kernel	54	29	315
$\text{Hier\_kernel}_{\text{sekv}}$	**50**	**26**	196
GMM	53	36	36

Tab.3: PER and MISS for different number of training sentences (for PLP features).

Number of sentences [-]	PER [%]	MISS [%]	Training time [min]
80	64	40	6,8
160	59	36	10
240	53	31	15,8

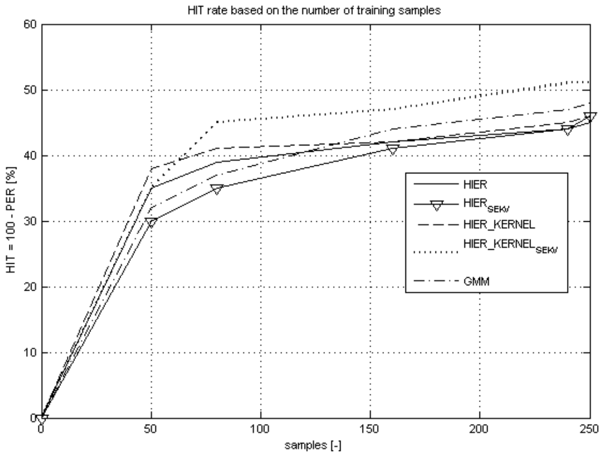


Fig. 5: Classifier precision (HIT) based on the number of input training samples.

to Contents, page 1

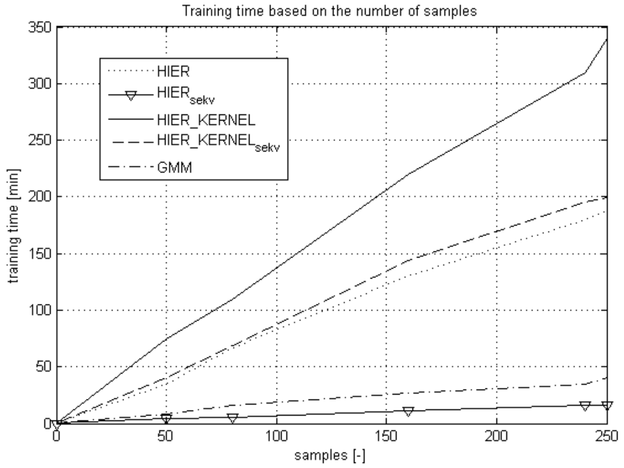


Fig. 6: Overall classifier training time.

7. Conclusion

This paper compared two state-of-art approaches to frame-based phoneme classification – generic and discriminative. Two state-of-art discriminative frame-based classifiers were presented (denoted as HIER and HIER_KERNEL) along with state-of-art generic classifier represented by the GMM frame-based classifier (denoted as GMM). Moreover, this paper had proposed two efficient training algorithms for discriminative frame-based phoneme classification (denoted with the SEKV suffix). For notation, all discriminative classifiers exploit a hierarchical tree root structure which is inducing tree root metric over the input group of phonemes. Both HIER and HIER_KERNEL classifiers had similar results on PER compared to the GMM classifier, but the results for metric MISS show the advantage of these classifiers (especially the implementation of hierarchical structure had shown to be a very effective). Both proposed training algorithms clearly outperforms all of the previous classifiers and showed possible future direction for frame-based phoneme classification. The results also showed superiority of the PLP features over the MFCC features.

Our future work will be focused on the implementation hierarchical tree root structure into the GMM classifiers and incorporation of long temporal content into the frame-based classifiers. The future effort will also aim on the classifiers evaluation within the KWS systems.

References

[1] PFEIFER, V.; BALIK, M.; MALY, J. Frame based phoneme classification using large margin and kernel methods. In *33rd International Conference on Telecommunications and Signal Processing, TSP 2010*, Baden, Austria, September 2010, p. 1-4. ISBN 978-963-88981-0-4.

[2] PFEIFER, V.; BALIK, M. Spotting techniques used in modern videoconference systems. In *31st International Conference on Telecommunications and Signal Processing, TSP 2008*, September 2008, pp. 55–58. ISBN 978-963-06-5487- 6.

[3] PFEIFER, V.; BALIK, M. Effective plagiarism detection system for advanced videoconference systems. In *32nd International Conference on Telecommunications and Signal Processing, TSP 2009*, Asszisztencia Szevezo kit. H-1136 Budapest. Hedegus Gyula u. 20, September 2009, pp. 101–106. ISBN 978-963-06-7716- 5.

[4] PFEIFER, V.; BALIK, M.; MICA, I.; PRUSA, Z. Phoneme confidence estimator based on gaussian mixture models. *Gests Internation Transactions on Communication and Signal Processing*, 2009. Vol. 13, No. 13, p. 1-8, ISSN 1738-9682.

[5] RFC 3611. *RTP* Control *Protocol Extended Reports*. Paris: The Internet Society, 2003. pp. 54

[6] RUSEK, Krzysztof; ORZECHOWSKI, Tomasz; DZIECH, Andrzej. LDA for Face Profile Detection. In *Communications in Computer and Information Science: Proceedings 4th International Conference Multimedia Communications, Services and Security Multimedia Communications, Services and Security 2011*. 1st ed. Berlin: Springer, 2011. pp. 144-148. ISBN 978-3-642-21511-7.

[7] DEKEL, O.; KESHET, J.; SINGEER, Y. *An online algorithm for hierarchical phoneme classification*. Springer, vol. 3361, 2005, pp. 146–158, January 2005. ISBN 3-540-24509-X.

[8] PFEIFER, V.; BALIK, M. Fonemovy klasikator zalozen na pravdepodobnostnich modelech. *Elektrorevue*, Vol. 64, p. 1-6, Prosinec 2009. Available at WWW: <http://www.elektrorevue.cz/cz/download/ramcova-hiearchicka-klasi-kace-fonem--ceskeho-jazyka>. ISSN 1213-1539.

[9] GREZL, F. *Trap-based probabilistic features for automatic speech recognition*. Brno, 2007. Ph.D. dissertation, Brno University of Technology.

[10] ARADILLA, G.; VEPA, J.; BOURLARD, H. *Using posterior-based features in template matching for speech recognition*. IDIAP research institute, Tech. Rep. IDIAP-PR 06-23, June 2006.

[11] KESHET, J. *Large margin algorithms for discriminative continuous speech recognition*. Hebrew, 2007. Ph.D. dissertation, Hebrew University.

[12] YE, J.; POVINELLI, R. J.; JOHNSON, M. T. Phoneme classification using naive bayes classifier in reconstructed phase space. In *Digital Signal Processing Workshop*, 2002 and the *2nd Signal Processing Education Workshop*. Proceedings of IEEE 10th, 2002. pp. 37–40. ISBN 0-7803-8116-5.

[13] GRANGIER, D.; BENGIO, S. *A discriminative decoder for the recognition of phoneme sequences*. IDIAP Research Institution, Tech. Rep., November 2005.

[14] KESHET, J.; BENGIO, S. Eds. *Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods*. Wiley, January 2009, vol. 1. pp. 268. ISBN 978-0470696835.

[15] KESHET, J.; GRANGIER, D.; BENGIO, S. Discriminative keyword spotting. In *Workshop on Non-Linear Speech Processing (NOLISP)*, 2007, p. 1-5. [Online]. Available at WWW: <http://david.grangier.info/pub/papers/ 2007/KeshetGrBe07.pdf>.

[16] GAROFOLO, J. S.; LAMEL, L. F.; FISHER, W. M.; FISCUS, J. G.; PALLETT, D. S.; DAHLGREN, N. L.; ZUE, V. Timit acoustic-phonetic continuous speech corpus. In *10th International Conference on Speech and Computer. Linguistic Data Consortium*, 1993, p. 4.

to Contents, page 1

[17] BISHOP, C. M. *Pattern recognition and Machine learning*, B. S. M. Jordan, J. Kleinberg, Ed. Springer, February 2006, Vol. 1. pp. 740. ISBN 978-0-387-31073-2.

[18] PSUTKA, J.; MULLER, L.; MATOUSEK, J.; RADOVA, V. *Mluvíme s počítačem česky*. ACADEMIA, October 2006. pp. 746. ISBN 80-200-1309-1.

[19] HERMANSKY, H.; ELLIS, D. P.; SHARMA, S. Tandem connectionist feature extraction for conventional hmm systems. In: *Proceedings of the conference ICASSP-2000*, 2000, p. 1-4. [Online]. Available at WWW: <ftp://ftp.icsi.berkeley.edu/pub/speech/papers/icassp00-nnhmm.pdf>. ISBN 0-7803-6293-4.

Vaclav PFEIFER was born in 1982. He received his M.Sc. from electronic and communications in 2006. His research interests include speech processing focused on the spoken term detection and reconstruction of the object trajectory based on the two rig cameras. He is author and co-author of many articles dealing with the spoken term detection and currently he is finalizing his thesis to complete Ph.D. studies.

Miroslav BALIK was born in 1973. He received the Ph.D. degree at the Faculty of Electrical Engineering and Communication of Brno University of Technology in 2003. In present time he focuses on noise reduction algorithms based on VOLA segmentation method. His research interests include “Digital Audio Signal Processing”, “Multimedia Systems and Services” and “Computer Systems”. He is the author of several digital systems for musical signal real-time processing. He is a senior member of IEEE.

Prime Research on Education (PRE)
ISSN: 2251-1253
Vol. 2(1), pp. 155-162, February 3rd, 2012
www. primejournal.org/PRE

to Contents, page 1

Full Length Research

The role of Emotional Intelligence (EI) and transformative leadership style of principals in high schools: An investigation

[1]Abbas Alavi shad and [2]Hamideh Vaezi

[1]University of Isfahan, Faculty of Administrative Sciences and Economic. E-mail: Alavishad@yahoo.om
[2]University of Kerman, Department of Educational Management. E-mail: Ha.vaezi@yahoo.com

Accepted 7th January, 2012

Considering the importance of human resource development goals of today's Schools and organizations, to address the patterns and techniques necessary for understanding and guiding staff is one of the most effective methods and guidance, familiarity with topics and transformative leadership styles and management practices is the transformative leadership style and management skills. The purpose of the present research as a descriptive-correlative research was to study the relationship between emotional intelligence and transformative leadership style of principals who work in secondary education. So five components include self-awareness, self-regulation, motivation, empathy and social skills and four styles of transformative leadership inspirational. Principals have high EI focus their efforts to create enthusiasm in their team with abundant energy and refer others to move forward. A principal's skill in the area of human relations, decision-making, control of subordinates and conflict resolution are indicators of transformative leadership traits and behaviors. Effective leaders will support and encourage staff to model behaviors promoting collegiality (Collegiality means a move away from an emphasis on all decision-making resting in the hands of one individual, towards a more shared and participative approach) and a professional working environment.

Key words: Emotional Intelligence (EI); transformative leadership styles; empathy; self-regulation; self-awareness; inspirational motivation; social skills; intellectual stimulation

INTRODUCTION

Most of the earlier theories have summarized intelligence as an educational ability and have focused on achievement talent. But, today there is a belief that some of the individual's non- intelligence characteristics such as being emotional, conscientious, having moral sensitivity and transformative leadership ability were mentioned as a separate dimension of intelligence. Many research studies dealing with defining intelligence, set forth that intelligence comes to existence through reciprocal effects of intelligence and non- intelligence characteristics, which include environmental skills which are to a great extent responsible for development and change in the present world.

In 1972 federal definition of intelligence was beyond the cognitive ability so that transformative leadership ability was mentioned as a separate and independent model of intelligence (Young Lee, 2006).

EI has been identified, through the popular press and some researchers as that critical element needed for effective transformative leadership.

In other Hand, transformative leadership, considered one of the most recent approaches to the investigation led to the country's few. Theory of transformative leadership style a theoretical framework is transformative in the world, which has been proposed by Burns (1978) and Bass (1985). Burns stated in 1990, stating that leaders can use behavioral characteristics of transformative leadership, the performance of their followers than expected.

The challenge with these standpoints is twofold, (1) the study of transformative leadership and what makes leaders effective has been found to be much more complicated than a single dimension like EI; and (2) organizations have incorporated many of these EI beliefs

to Contents, page 1

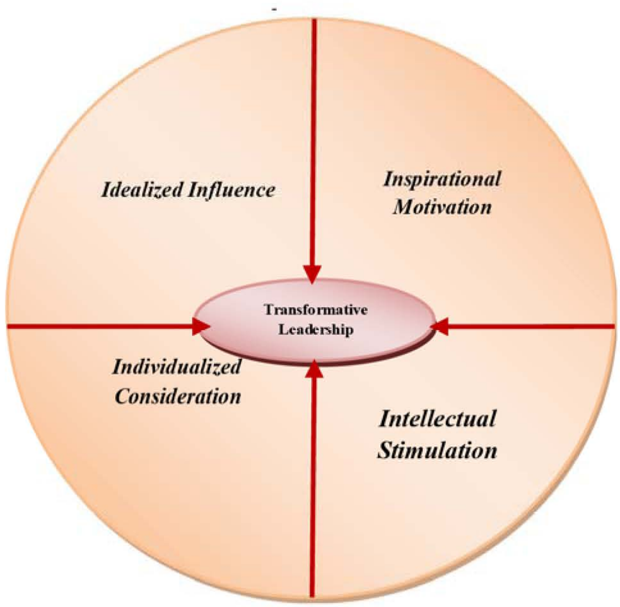


Figure 1: Model of transformational leadership

into their work systems and performance expectations without researching what some authors claim is true and achievable (Salovey, 2003). The study of transformative leadership, its effectiveness and its impact on organizational performance is a key interest to Human Resource Development (HRD) scholars. On this basis, in this paper we review the literature on the EI and transformative leadership Style and We Survey the relationship between them and principals in high schools

Transformative Leadership

Accomplishing the kinds of changes needed to integrate EI into secondary schools requires transformative leadership: leadership that is willing to realign structures and relationships to achieve genuine and sustainable change. Although there are more elements of transformational leadership than we can elaborate here, we can describe some key aspects derived from education research and stories of successes (Bencivenga and Elias, 2003; Devaney et al., 2006; Elias and Arnold, 2006; Elias, Arnold, and Hussey, 2003). For bringing major changes, transformational leaders must Consider the following four styles (figure 1):

Inspirational Motivation: The foundation of transformational leadership is the promotion of consistent vision, mission, and a set of values to the members. Their vision is so compelling that they know what they want from every interaction. Transformational leaders guide followers by providing them with a sense of meaning and challenge. They work enthusiastically and optimistically to foster the spirit of teamwork and commitment. ***Intellectual Stimulation*** :Such leaders encourage their followers to be innovative and creative. They encourage new ideas from their followers and never criticize them publicly for the mistakes committed by them. The leaders focus on the "what" in problems and do not focus on the blaming part of it. They have no hesitation in discarding an old practice set by them if it is found ineffective.
Idealized Influence :They believe in the philosophy that a leader can influence followers only when he practices what he preaches. The leaders act as role models that followers seek to emulate. Such leaders always win the trust and respect of their followers through their action. They typically place their followers needs over their own, sacrifice their personal gains for them, ad demonstrate high standards of ethical conduct. The use of power by such leaders is aimed at influencing them to strive for the common goals of the organization.
Individualized Consideration: Leaders act as mentors to their followers and reward them for creativity and innovation. The followers are treated differently according to their talents and knowledge. They are empowered to make decisions and are always provided with the needed support to implement their decisions. Overall, research shows that the four factors of transformative leadership inspirational the two styles - inspirational motivation and intellectual stimulation are most important factors related to the field of education.

DISCATION

Studies Alon and Higgins (2006) show that the fact that emotional intelligence (EQ), analytical intelligence (IQ), and transformative leadership behaviors are moderated by cultural intelligence (CQ) in the formation of global transformative leadership success.

In their study of the relationship between emotion and Transformative leadership, Gardner, Fischer, and Hunt (2009) reviewed the literature of emotional labor and authentic transformative leadership and identified three categories of leader emotional displays: surface acting, deep acting and genuine emotions. "The consistency of expressed leader emotions with affective display rules, together with the type of display chosen, combines to impact the leader's felt authenticity, the favorability of follower impressions, and the perceived authenticity of the leader by the followers. They also explored the influence on leader emotional labor of contextual dimensions of the environment, including the omnibus (national and organizational culture, industry and occupation, organizational structure, time) and discrete (situational) context.

Goleman and his colleagues (2003) examined the relationship between EI and effective performance, especially in leaders. They observed to what degree emotional intelligence manifests itself in the work place. Alavishad (2010) research was designed to determine which personal capabilities drove outstanding performance. He grouped the skills into three categories; skills, cognitive skills, and competencies demonstrating

EI. His data revealed dramatic results. Cherniss (2010) states, “My analysis showed that EI played an increasingly important role at the highest levels of the company”.

McKee (2002) posited that leaders use EI to develop relationships that are in-sync with their organization by forming “emotional bonds that help them stay focused even amid profound change and uncertainty.” Essentially, the Teachers of the future will need to be attuned to the big picture, and be able to think conceptually as they transform the organization through people and teams. They will also need to possess strong interpersonal skills, be able to get along with others, and exercise high levels of intelligence and energy.

Fullan (2009) emphasized that “emotionally intelligent leaders are aware of their own emotional makeup, are sensitive and inspiring to others, and are able to deal with day-to-day problems as they work on more fundamental changes in the culture of the organization”.

School leaders are faced with an abundance of issues when they assume a transformative leadership position, second only to high expectations for systemic and transformational change in the school system.

Meyerson, Orr and Cohen (2007) presented in their report, Preparing School Leaders for a Changing World, key components necessary for exemplary principal preparation programs. The recommendations proposed in these reports are valid, but equally important is the balance of training in the area of EI for an educational leader’s success in becoming a change agent for the improvement of instruction. As defined by Goleman (2004), EI is the ability to lead, recognizing four emotional areas: self-awareness, self-management, social awareness, and relationship management, each having specific characteristics. These four cluster areas focus on identified traits, behaviors and characteristics of successful leaders. Research has identified additional areas including organizational and management skills, shared values and beliefs, collegiality, and staff building. In each of these areas EI is a common theme.

Marzano (2003) highlights three principles for effective leaders. The first revolves around the principal functioning as a strong cohesive force; the second is to provide strong guidance while demonstrating respect; and the third principle is characterized by specific behaviors which enhance interpersonal relationships. Principle three further establishes three characteristics of importance: optimism, honesty, and consideration. Optimism increases teachers’ self-esteem and motivation. “Honesty is characterized by truthfulness and consistency between words and actions”.

Hausman, Crow, and Sperry (2005) contend, “Their actions are congruent with their values.” The authors continue stressing the need for the leader to understand their needs and emotions as well as their strengths and limitations. “The ideal Principals must focus intensely on

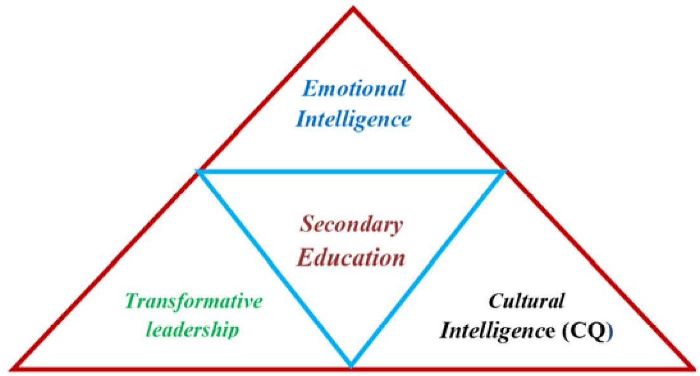


Figure 2: Interaction EI ,Transformative leadership and CQ Education
Sourse: (Wolff, Pescosohdob and Druskatc, 2008).

their interpersonal skills, capacity to read and adjust to the environment, and the ability to understand and cope with far ranging issues. They must be politically astute, prepared to adjust their Transformative leadership styles, and ethically grounded” (Hausman, Crow and Sperry, 2005). Figure 2

Learning experiences for Principals cannot just reinforce old “platitudes” of being effective, but must encourage Principals to question their practices and attempt (Youn Lee, and Olszewski-Kubilius, 2006). At times transformative leadership is viewed as a mysterious and elusive concept. The challenge is for individuals to look inward to achieve effective transformative leadership (Chopra, 2002).

As Dewey advocated the teaching of the “whole child” for maximum gains, so should programs for transformative leadership include the social, emotional, intellectual and physical components. It is through the combination of these focused areas that transference of meaningful change will take place in our schools. Strong transformative leadership development processes are focused on emotional and intellectual learning and they build on active participatory work: action learning and coaching, where people used what they’re learning to diagnose and solve real problems in their organizations.(Goleman, Boyatzis, and McKee, 2002)So successful schools need educational leaders who have the abilities to facilitate sustained and lasting change. As Fullan (2009) reiterated, “these new educational leaders will need to have a strong sense of moral purpose for direction and great EI as they build relationships”. In effect, Principals will not only need strong intellectual skills to be a great influence on the school culture, but they must be able to influence and understand relationships and the feelings and emotions of those they serve and lead) (Moss and Ngu, 2006). The research basis purpose, then, was to investigate the impact of EI on Transformative leadership styles of Principals who

to Contents, page 1

Table 1: Correlation coefficient between Emotional Intelligence and principal's transformative leadership styles

Transformative leadership Styles	Emotional Intelligence		
	n	r	Sig
Inspirational Motivation style	42	0.714	0.000
Intellectual Stimulation style	42	-.0719	0.000

Table 2: Correlation coefficient between Self-awareness and principal's transformative leadership

Transformative leadership Styles	Self-awareness		
	n	r	Sig
Inspirational Motivation style	42	0.296	0.057
Intellectual Stimulation style	42	-0.242	0.123

work in Secondary Education in Iran.

METHODOLOGY

Population and sampling: Research method was descriptive- correlative. Statistical population included all teachers of high schools of Khoramabad, a small city in Iran. The sample included 42 high school and 252 teachers that were chosen by proportional stratified sampling were selected as the sample of this study.

The research objectives

1. Determine the relationship between Self-awareness and transformative leadership styles of high school principals in Khoramabad city.
2. Determine the relationship between between Self-regulation and transformative leadership styles of high school principals in Khoramabad city.
3. Determine the relationship between motivation and Inspirational Motivation transformative leadership styles of high school principals in Khoramabad city.
4. There is a significant relationship between empathy and transformative leadership styles of high school principals in Khoramabad city.
5. There is a significant relationship between Social skills and transformative leadership styles of high school principals in Khoramabad city.
6. There is a significant relationship between EI and its components according to age and gender

Instruments: Research instruments were two questionnaires of transformative leadership styles profile (TLSP) and EI of shrink.

Shrink's EI questionnaire: This questionnaire has 33 items, which has been devised Shrink on the basis of Goleman 's theory. The reliability of this questionnaire was reported to be 0. 82. The findings of the present study also showed that EI scale had a very high reliability.

Transformative leadership styles questionnaire: This questionnaire measures transformative leadership styles in 32 items. A pilot study of the questionnaire with 30 managers revealed an Alpha Cronbach coefficient of 0.86 for this scale.

Data Analysis

General Hypothesis

There is a significant relationship between EI and the transformative leadership styles of high school principals in Khoramabad city. The Pearson correlation coefficient was used to examine this hypothesis. The result is reported in table 1.

As it has been shown in table 1, correlation coefficient between EI and Inspirational Motivation Transformative leadership style, $r= 0.714$ is significant at level $p<0.000$. Accordingly, there is a significant relationship between EI and Inspirational motivation transformative leadership style. It means that the teachers with high EI apply Inspirational Motivation Transformative leadership style more.

The result also shows that there is a significant relationship between EI and Intellectual Stimulation transformative leadership style with $r= -0.719$ which is significant at level $p<0.000$. Accordingly, there is a negative relationship between EI and Intellectual Stimulation transformative leadership style. It means that the teachers with high EI apply Intellectual Stimulation transformative leadership style less. According to the obtained results, "null hypotheses" as a clue of no relationship between EI and the teachers' transformative leadership style cannot be confirmed.

Hypothesis 1

There is a significant relationship between Self-awareness and transformative leadership styles. Pearson correlation coefficient has been used to examine this hypothesis. The results are reported in table 2.

As it has been shown in table 2, correlation coefficient between Self-awareness and Inspirational Motivation transformative leadership style, $r= 0.296$ is not significant at level $p< 0/000$. Accordingly, there is not a significant

Table 3: Correlation coefficient between Self-regulation and principal's transformative leadership

Transformative leadership Styles	Self-regulation		
	n	r	Sig
Inspirational Motivation style	42	0.420	0.006
Intellectual Stimulation style	42	-0.472	0.002

Table 4: Correlation coefficient between Motivation and principal's transformative leadership style.

Transformative leadership Styles	Self-regulation		
	n	r	Sig
Inspirational Motivation style	42	0.515	0.000
Intellectual Stimulation style	42	-0.481	0.001

relationship between Self-awareness and inspirational motivation transformative leadership style.

Also table 2 shows that there is not a significant relationship between Self-awareness and intellectual Stimulation transformative leadership style. The Pearson correlation coefficient between Self-awareness and Intellectual Stimulation transformative leadership style (r= -0.242) is significant at level p<0/000. Accordingly, there is a negative relationship between Self-awareness and Intellectual Stimulation transformative leadership style. Thus "null hypotheses" as a clue of no relationship between Self-awareness the teachers' Transformative leadership style cannot be confirmed but instead the "research hypothesis" is confirmed. It means that there is not a significant relationship between Self-awareness and Inspirational Motivation styles and also Intellectual Stimulation styles.

Hypothesis 2

There is a significant relationship between Self-regulation and transformative leadership styles. Pearson correlation coefficient has been used to examine this hypothesis. The results have been given in table 3.

As it has been shown in table 3, the correlation coefficient (r= 0.420) between Self-regulation and inspirational motivation transformative leadership style is significant at level p< 0/006. Accordingly, there is a significant relationship between Self-regulation and inspirational motivation transformative leadership style. It means that teachers with high Self-regulation apply Inspirational Motivation transformative leadership style more. In addition the relationship obtained from table 3 show that there is a significant relationship between Self-regulation and Intellectual Stimulation transformative leadership style. Pearson correlation coefficient between Self-regulation and Intellectual Stimulation transformative leadership style (r= -0.472) is significant at level p<0.002. Accordingly, there is a negative relationship between Self-regulation and Intellectual Stimulation transformative leadership style. It means that the managers with high Self-regulation apply Intellectual Stimulation transformative leadership style less. According to the obtained results, "null hypotheses" as a clue of no relationship between Self-regulation the managers' transformative leadership style cannot be confirmed but instead "research hypothesis" is confirmed. It means that there is a significant relationship between Self-regulation and transformative leadership styles.

Hypothesis 3

There is a significant relationship between motivation and inspirational motivation transformative leadership styles. Pearson correlation coefficient has been used to examine this hypothesis. The results have been given in table 4.

As it has been shown in table 4, correlation coefficient between motivation and inspirational motivation transformative leadership style (r= 0.515) is significant at level p<0.000. Accordingly there is a significant relationship between motivation and inspirational motivation transformative leadership style. It means that teachers with high motivation apply inspirational motivation transformative leadership style more.

Also the relationship obtained from table 4 shows a significant relationship between motivation and intellectual stimulation transformative leadership style. Pearson correlation coefficient between motivation and Intellectual Stimulation transformative leadership style (r= -0.481) is significant at level p<0/001. Accordingly there is a negative relationship between motivation and Intellectual Stimulation transformative leadership style. It means that teachers with high motivation apply intellectual stimulation transformative leadership style less. According to the obtained results, "null hypotheses" as a clue of no relationship between motivation and Teachers' transformative leadership style cannot be confirmed but instead "research hypothesis" is confirmed. It means that there is a significant relationship between motivation and transformative leadership styles.

Hypothesis 4

There is a significant relationship between empathy and transformative leadership styles. Pearson correlation coefficient has been used to examine this hypothesis.

to Contents, page 1

Table 5: Correlation coefficient between Social skills and principal's Transformative leadership style

Transformative leadership Styles	Social skills		
	n	r	Sig
Inspirational Motivation style	42	0.367	0.017
Intellectual Stimulation style	42	-0.355	0.021

Table 6: Correlation coefficient between significant relationship and principal's transformative leadership style

Transformative leadership Styles	significant relationship		
	n	r	Sig
Inspirational Motivation style	42	0.459	0.002
Intellectual Stimulation style	42	-.557	0.000

The results have been given in table 5.

As it has been shown in table 5, the correlation coefficient between empathy and inspirational motivation transformative leadership style (r= 0.367) is significant at level p< 0.017. Accordingly, there is a significant relationship between Empathy and inspirational motivation transformative leadership style. It means that Teachers with high empathy apply inspirational motivation transformative leadership style more. The relationship obtained from table 5 also shows that there is a significant relationship between empathy and intellectual stimulation transformative leadership style. Pearson correlation coefficient between Empathy and intellectual stimulation transformative leadership style (r= -0.355) is significant at level p<0.021. Accordingly there is a negative relationship between empathy and intellectual stimulation transformative leadership style. It means that Teachers with high empathy apply intellectual stimulation transformative leadership style less. According to the obtained results, “null hypotheses” as a clue of no relationship between empathy and Teachers' Transformative leadership style cannot be confirmed but instead “research hypothesis” is confirmed. It means that there is a significant relationship between empathy and inspirational motivation styles and intellectual stimulation transformative leadership styles.

Hypothesis 5

There is a significant relationship between Social skills and Transformative leadership styles. Pearson correlation coefficient has been used to examine this hypothesis. The results have been given in table 6.

As it has been shown in table 6, correlation coefficient between social skills and Inspirational Motivation Transformative leadership style (r= 0.459) is significant at level p< 0.002. Accordingly there is a significant relationship between Social skills and inspirational motivation transformative leadership style. It means that Teachers with high social skills apply inspirational motivation transformative leadership style more.

The relationship obtained from table 6 also shows that there is a significant relationship between social skills and intellectual stimulation transformative leadership style. Pearson correlation coefficient between social skills and intellectual stimulation transformative leadership style (r= -0.557) is significant at level p<0/000. Accordingly there is a negative relationship between social skills and intellectual stimulation transformative leadership style. It means that teachers with high EI apply intellectual stimulation transformative leadership style less.

According to the obtained results, “null hypotheses” as a clue of no relationship between social skills and teachers' transformative leadership style cannot be confirmed but instead the “research hypothesis” is confirmed. It means that there is a significant relationship between social skills and inspirational motivation and close transformative leadership styles.

Hypothesis 6

There is a significant relationship between EI and its components according to age and gender.

(a) The relationship between Teachers' gender and EI and its components At test was used to compare the differences between EI and its dimensions in Teachers according to gender. The results have been reported in table 7.

The result of t Test showed that there was no significant relationship between Teachers' EI and its components (t=0.89, sig=0.78, P,0.05). Moreover, no relationship between gender and EI's components was observed. Accordingly, the “null hypothesis” is confirmed.

(b) The relationship between Teachers' age and EI and its components To find out the relationship between Teachers' age and EI and its dimensions, a One Way ANOVA was used. The results have been reported in table 8.

The results of ANOVA reported in table 8 showed that there was no relationship between Teachers' EI and age. F=0.648, significance=0.528 at p<0.05 is not significant. Just in Self-awareness the obtained F is significant (F=3.42, significance=0.043, p<0.05). Other components

Table 7: The results of t test for differences between EI and its dimensions in teachers according to gender

Variables		n	mean	SD	t	df	sig
EI	Female	22	124.9	.89	.0.89	40	0.78
	Male	20	122.2	1.2			
Self-awareness	Female	22	31.45	1.9	10.9	40	0.69
	Male	20	32.65	085			
Self-regulation	Female	22	26.14	-1.97	-1.97	40	0.48
	Male	20	24.00				
Motivation	Female	22	23.95	.26	.26	40	0.46
	Male	20	23.70				
Empathy	Female	22	23.14	.80	.80	40	0.89
	Male	20	22.25	3.30			

Table 8: The results of one way ANOVA for the relationship between teachers' age and EI

Variables	Sum of squares	df	mean of squares	F	Sig	
EI	Between Groups Within groups total	125.15 3764.19 3889.34	2 39 41	62.57 96.52	0.648	0.528
Self-awareness	Between Groups Within groups total	77.43 441.54 518.97	2 39 41	38.7 11.32	3.42	0.043
Self-regulation	Between Groups Within groups total	4.36 536.04 540	2 39 41	2.18 13.75	0.16	0.85
Motivation	Between Groups Within groups total	13.64 400.18 413.83	2 39 41	6.82 10.26	0.66	0.52
Empathy	Between Groups Within groups total	45.70 468.86 514.57	2 39 41	22.85 12.02	1.90	0.16
Social skills	Between Groups Within groups total	10.90 317.10 328.00	2 39 41	5.45 8.13	0.67	0.517

of EI and teachers' age were not significantly different. Accordingly, the "null hypothesis" is confirmed and the "research hypothesis" is not confirmed.

CONCLUSION

Based on the results of general hypothesis, it can be concluded that there is a significant relationship between EI and Inspirational Motivation Transformative leadership style. It means that teachers with high EI apply Inspirational Motivation Transformative leadership style more and try hard to make enthusiasm among their team members, also encourage others strongly to make them sustain trying. The findings suggest that among organizational duties there should be performing plans related to the EI skills. In addition, in the process of appointing managers and employees, EI should be considered as one of the criteria. The high school Teachers with emotional intelligence can achieve the various goals like removing the obstacles, solving the dissensions and also improving the education outcome in school.

Based on the results of six hypotheses, it can be concluded that the results related to differences of EI and its componens among high school teachers and their gender are in agreement with Bryan(2007), Mayer (2008), Alavishad (2010). Studies. Golman (2002) also believes that the gender differences have no influence on EI. Also based on the obtained results" null hypothesis" as a clue of no relationship between EI and the age of high school Teachers was confirmed. Besides "research hypothesis" was not confirmed too. The result is in agreement with studies carried out by, Alavishad (2010).As the results showed, EI characteristics have been recognized as positive attributes in effective leaders. The characteristics are attributes associated with success and the frequency of the "emotional" trait was strong, as cited by Kouzes and Posner, (2002) Maxwell, (2004) and Sergiovanni (2002). The question remains, how do we prepare and mentor future administrators for success in leading transformational change in our school system? In order

to Contents, page 1

for collaboration, response and mobilization to occur, self-reflection on the part of the leader is the starting point for successful relationships within the school community. To promote success for all students, leaders must become acquainted with the areas related to EI and the competencies necessary to be successful. Educational Transformative leadership programs should include EI theory as a component for reform. Programs have been focusing on the development of course content; the time has come to embrace the research on EI and provide a balanced approach.

REFERENCES

Alon I, Higgins J (2006). Global Transformative leadership success through emotional and cultural intelligences Business Horizons, 48: 501-512.

Bryan SP (2007). Emotional Intelligence and Intrapersonal Conversations. Published in www.eiconsortium.org. The Consortium for Research on Emotional Intelligence in Organizations – Issues in Emotional Intelligence. pp. 1-9.

Cherniss C (2010). Emotional intelligence: What is it and why it matters. Paper presented at the annual meeting of the Society for Industrial Organizational Psychology. Rutgers University, Graduate School of Applied and Professional Psychology. Web site: www.eiconsortium

Chopra D (2002). The soul of transformative leadership . [Electronic version]. School Administrator, 59(8), pp1-10

Cohen C (2007). Preparing School Leaders for a Changing World: Lessons from Exemplary Transformative leadership Development Programs. Stanford, CA: Stanford University, Stanford Educational transformative leadership Institute.

Fullan M (2009). Teachers as leaders in a culture of change. ducational Transformative leadership , 5, 16-20.

Gardner WL, Fischer D, Hunt J (2009). Emotional labor and Transformative leadership: A threat to authentic city? The Transformative leadership quarterly, 20: 466-482.

Goleman D (2003). Emotional intelligence: Why it can matter more than IQ. New York: Bantam Books.

Goleman D (2004), "What makes a leader?" Harvard Business Review, 76: 93- 104

Goleman D, Boyatzis R, McKee A (2002). Primal Transformative leadership : Realizing the power of emotional intelligence. Boston: Harvard Business School Press.

Hausman C, Crow GM, Sperry D (2005,). Portrait of the "ideal principal": Context and self. [Electronic version]. NASSP Bulletin, 84(617): 5-14.

Alavishad Lila (2010). Investigated the Relationship between Emotional Intelligence and the Heads Secondary Schools in Tehran. MSc thesis in Shahid Beheshti University, Faculty of Educational Sciences and Psychology.

Kouzes JM, Posner BZ (2002). The Transformative leadership challenge. San Francisco, CA: Jossey-Bass Inc.

Mayer JD (2008). Emotion, Intelligence and Emotional Intelligence. Handbook of affect and Social cognition (pp.30-49). Mahwah, NJ: Lawrence and Erlbaum Association.

Marzano R (2003). What works in schools. Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).

Moss S, Ngu S (2006). The Effect of Follower Regulatory Focus and Extraversion on transformative leadership Behavior: The Role of Emotional, Intel. J. Individ. Diff., 27: 93-107.

Maxwell J (2004). The 21 indispensable qualities of a leader. Tennessee: Thomas Nelson Inc.

Salovey P (2003). Emotional intelligence. Imagination, Cognition, and Personality, 9: 185-211.

Sergiovanni TJ (2002). Moral Transformative leadership : Getting to the heart of school improvement. San Francisco, CA: Jossey-Bass Inc.

Wolff BS, Pescosohdob TA, Druskatc UV (2008). Emotional Intelligence as the basis of Transformative leadership emergence in self- managing team. J. Leader Quar., 13: 505-522. DOI: 10.1016/S1048-9843(02)00141-8.

Youn LS, Olszewski-Kubilius P (2006). The Emotional Intelligence, moral judgment and Transformative leadership of academically gifted adolescents. J. Edu. Gif, 30: 29-67.

to Contents, page 1

PRIME Journals
Prime Reseach on Education (PRE)
ISSN: 2251-1261 Vol. 2(1), pp. 163-168, February 3rd, 2012
www. primejournal. org/PRE

Full Length Research

Predictive validity of body parameters on academic performance of Nigerian primary school pupils in mathematics scholastic aptitude test

***EB Kolawole[1] and DO Udeh[2]**

[1]*Institute of Education,Faculty of Education,Ekiti State University, Ado-Ekiti, Nigeria. E-mail: ebkolawole@yahoo.com
[2]Department of Mathematics and Computer Sciences, Government Science College, Ayede-Ekiti, Nigeria.
E-mail: udohprof@yahoo.com

Accepted 1st February, 2012

The study is a predictive analysis to determine the relationship and the predictive strength between the body parameters of head circumference, body weight, body height, body Mass Index (BMI), palm length, feet length and measure of pupils' academic performance in mathematics scholastic academic test. The study employed survey research design. The population of the study comprised all primary school pupils in Ekiti State, Nigeria. Nine hundred pupils were sampled from 9 local government areas of the State, using multi–stage random sampling technique. An adapted test instrument tagged: "Pupils Mathematics Scholastic Test" (PMSAT) was used to elicit pupils' academic performance in Mathematics aptitude test. Appropriate instruments were used for the measurement of the body parameters. The PMSAT was certified to have face validity and predictive validity of 0.88 and a reliability coefficient of 0.85 which was ensured through split–half method and corrected with Spearman Brown prophecy formular. The measuring instruments for the body parameters have high validity and reliability since they are instruments used in physical sciences which are very objective and accurate. The researchers administered the PMSAT test on the pupils and measure their body parameters. Data collected were analyzed using descriptive and inferential statistics at $\alpha = 0.05$ level of significance. The findings revealed that there is a positive significant relationship between body parameters and pupils' academic performance in mathematics scholastic aptitude test. Body height is the best predictor of pupils' academic performance in mathematics scholastic aptitude test, while body mass index (BMI) is the worst predictor of pupils' academic performance in mathematics scholastic aptitude test. Head circumference, body weight, palm length and feet length have no significant effect on pupils' academic performance in mathematics scholastic aptitude test. We recommend that pregnant women should be placed on a health balance diet throughout the period of pregnancy and more especially during the first trimester. At birth, a child's body parameters should be measured and if there is an indication of a stunted growth, then growth hormone (GH) could be applied so that the child could catch up growth.

Keywords: Body parameters, predictive validity,primary school, achievement test, mathematics, academic performance

INTRODUCTION

Galton (1869) cited in Miller (1982) in his research to gather data about the dimensions of a man's mind he discovered that physical body measurements or the anthropometric measures of stature, weight, length, the size of the a man's skill, and so on, is closely related to mental measurements. Several researchers have studied either one of the body parameters such as the bodyweight, body height, head circumference and so on, and found that they have relationship with academic performance. Schumacher, and Desimone (2005) found that weight gain and being overweight and Obese each reduces academic performance by a significant amount that is typically statistical. Similar findings were made by Ashlesha et al., (2004) in a national study of childhood overweight and academic performance of kindergartens and first graders, asserted that overweight children have significantly lower Maths and Reading test scores compared with non-overweight children in kindergarten. Both groups were gaining similarly on maths and reading tests scores, resulting in significantly lower test scores

to Contents, page 1

among overweight children at the end of grade one.

Sergeant and Blanch-flower (1994) quoted by Sabia (2007) found that females who were Obese at age 16 had lower reading and math test scores later in life than those who were not obese at the same age. Moreover, Crosnoe and Muller (2004) cited in Sabia (2007) showed that adolescents in the 85th or higher percentile of the Body Mass Index (BMI) distribution for their age-gender group have lower mean Grade Points Average (GPAs) than those in the lower 85th percentile of the distribution. They find out that GPAs were even lower for obese adolescents in schools with higher rates of romantic relationships and lower average body size among students. Cawley and Spiess (2008), emphasized that body weight, do negatively affect academic performance and skill attainment in early childhood.

In a similar study on primary school children in Li, (1995) found that severely obese children had significantly lower IQ on math scores than the controls. He further averred that, IQ is a measure of ability and consequently it would be likely to affect school performance, but not an indicator of academic achievement. In another development, Mo-Suwan et al., (1999) examined the relationship between overweight status and academic performance from grades 3 to 6 and 7 to 9, using data from Thailand, concluded that being overweight during adolescence (grade 7 to 9) was associated with poor school performance, whereas such an association did not exist in children (grades 3-to-6). According to David et al., (2007), in Wikipedia (2010) Obese children had statistically lower course grades in Reading, Maths, Writing and Science measures, as well as weight increases academic performance decreases. The reason to this is that obesity affects children's psychological outcomes such as low self-esteem, low self-image, low self-efficacy and depression, resulting from overweight concerns. These psychological consequences of obesity combine with continuing discrimination; result in reduced life chances and poor academic performance with potentially even more serious adverse social outcomes in the long term (French et al., 1995; and 2003; Strauss, 2000). Kim et al, (2003) stated that the height of children positively correlates with their academic performance. He further asserted that other quantities such as intelligence, social class status and class mobility also show a positive correlation with height. This implies that taller children usually have higher academic performance than short children. Similar conclusion was reached by Ogunshola (2009) who found that height of academically poor pupils were lower than the height of academically good ones. Joel cited in Encyclopaedia (2010) stated that height is a biomarker of nutritional status or general mental and physical health during development. Case and Pexson (2008b) noted that the pre-natal environment and nutrition during childhood determine both body height and cognitive ability. Hence intrauterine malnutrition during pregnancy affects brain development, and relative deficiency of the Growth hormone (GH) - Like growth Factor (LGF) -1 axis (Winick and Rosso, 1969; Harvey et al; 1982; Georgieft, 1998). Pasquino et al; (2001) further asserted that children who lack GH reach an average mature height of only 4 feet, 4 inches. Although, when treated early with injection of GH, such children show a catch-up growth. In addition, Yvonne et al., (2004) in a study of 2 year of GH treatment in short children, born SGA, found a significant increase in total IQ score, in social acceptance scores, and in general self-worth scores. From the GH treatment result, it revealed by implication that short children have lower IQ score, poor psycho social functioning and poor self-concept. Similarly several studies have equally shown a positive correlation between intelligence and height (Walker et al., 2000; Pearce et al., 2005). Similar association have been found in early and late childhood and adulthood in both developed and developing countries and associations persisted after controlling for social class and parental education (Encyclopaedia, 2010). Case and Paxson (2008b) admitted that height positively associates with cognitive abilities (that is. intelligence). They further opined that height is positively correlated with cognitive abilities already at age 3 and throughout childhood which affect academic performance. They argued further that taller workers have an average higher wages because they are more intelligent.

Moreover, Venom quoted by Wikipedia (2009) in an investigation, showed that external head size measures of the head length, head circumference correlated with I .Q scores and school performance. According to Fauziah (2010) head growth in foetal life and infancy is associated with later intelligence. He added that brain growth in early life may be important in determining not only the level of peak cognitive function attained but also whether such functions are preserved in old age. He further asserted that older people with a larger head circumference tend to perform better on tests of cognitive function and may have reduced risks of cognitive decline and Acclaimer's disease.

He further reported that prenatal head growth and head growth during infancy remained significantly predictor for later I.Q., with full-scale I.Q. increasing on average of 1.56 points for each 1-SD increase in head growth. They concluded that head circumference at birth was no longer associated with IQ at age 8. This implies that the I.Q of a child is dependent on the head size at birth, early childhood and not the head size at later childhood, adolescence or adulthood. Consequently, Conner et al., (2001) and Schonfeld et al., (2001) observed that small head is an indication that the brain has not developed fully. According to Ophandodor(2003) in Ogunshola (2009) the correlation between small head circumference, or microcephaly and mental retardation, indicates that head circumference was an excellent predictor of I.Q. and academic performance. Adima (1989) referred to small head circumference or microcephaly as 'coconut head'. He further affirmed that these heads are usually small, due to an abnormally small brain and small skull, which made them profoundly mentally retarded.

Moreover, Stoch et al., (1982) added that learning and behaviour are seriously affected and an improved diet did not result in catch-up in head size. This may be as a result of malnutrition after the first trimester which causes foetuses to be born underweight and had small heads (Stein et al., 1975). Morgane et al., (1993) admitted that the poorer the mother's diet, the greater the loss in brain weight.

Zhang, (2010), citied in The Nation Newspaper (2010, July 5) submitted that malnutrition early in life appears to diminish brain function in older adulthood. They observed across the world, that 178 million children under age five are stunted or short in stature due to hunger, infection or both. The survey equally included a screening test for cognitive impairment, measurements of arms and lower legs, indicated childhood malnutrition or infection which put a question on childhood hunger. Galler,et al, (1984, 1990) suggested that the malnutrition have probably interfered with myelinization, causing a permanent loss in brain weight. As a result these children score low on intelligence test, show poor fine motor co-ordination, and have difficulty paying attention. And Owuamanam and Owuamanam (2004) confirmed that poor coordination will invariably affect the child's academic skills in the areas of writing, reading and handicraft. However, this study examined whether body parameters of head circumference, body height, body weight, body mass index (BMI), palm length and feet length have any relationship and effect on pupils' academic performance in Mathematics achievement test.

Statement of the problem

The problem of the study is the poor academic performance among Nigerian pupils and students in mathematics, in both internal and external examinations and the factor(s) responsible. To this end, this study is designed to investigate whether the body parameters of the pupils are responsible for the poor academic performance in mathematics scholastic aptitude test.

Purpose of the study

The purpose of the study is to determine whether the body parameters affect the academic performance of Nigerian primary school pupils. It seeks to find out the effect of each of the body parameters on the academic performance in Mathematics scholastic aptitude test, in order to determine the best and the worst predictors of the academic performance in Mathematics scholastic aptitude test.

Research questions

In order to solve this problem the following questions were raised:

1. Is there any relationship between the body parameters and pupils' academic performance in mathematics scholastic aptitude test?
2. Is there any effect of body parameters on pupils' academic performance in mathematics scholastic aptitude test?
3. Is there any predictive validity of pupils' academic performance in mathematics scholastic aptitude test on body parameters?

Research hypotheses

Based on the generated questions the following hypotheses were postulated:

1. There is no significant relationship between the body parameters and pupils' academic performance in mathematics scholastic aptitude test.
2. There is no significant effect of body parameters on pupils' academic performance in mathematics scholastic aptitude test.
3. There is no significant predictive validity of pupils' academic performance in mathematics scholastic aptitude test on body parameters.

METHODOLOGY

This study adopted the descriptive research of the correlation design. The population of the study comprised of 155,290 pupils of both public and private primary school pupils in Ekiti State, Nigeria. The sample comprised of 900 pupils of primary V classes for private primary schools and primary VI classes for public schools, within the age range of 9-to-13 and inclusive, which were sampled using multi-stage random sampling techniques from the three senatorial districts of Ekiti State, Nigeria.

The test instrument used to elicit the pupils' achievement in mathematics was adapted from Kolawole(2011) tagged: Pupils Mathematics Scholastic Aptitude Test (PMSAT), containing 50 multiple choice objective item type, with 5 options (A-E)The instruments used to sample the body parameters of body height, body weight, head circumference, palm length and feet length are: two meter rules of 1 meter each were joined together and fixed on the smooth wall, was used to measure the pupils' body height in metre (m), Harson weighing balance of model H89 was used to measure the pupils' body weight in kilogram (kg). For head circumference, palm length and feet length, measuring tape was used to measure them in centimetre (cm). All the body parameters were recorded on a Proforma titled, "Anthropometric Measures of Primary School Pupils" (AMPSP). The test instrument PMSAT was certified to have face validity and content by an expert. The reliability coefficient of 0.85 for PMSAT was obtained through the split-half method and Spearman Brown prophecy formula. The meter rules, weighing balance and measuring tape are objective and accurate measuring instrument used in the physical sciences, hence, it is reliable and valid. The instruments were administrated by the researchers. Data collected for the study were analyzed using descriptive statistics of mean, standard

to Contents, page 1

Table 1: Summary of Correlation Analysis and ANOVA between the Body Parameters and PMSAT

Sources of variation	N	Se	R^2	R_{cal}	R_{tab}	$F_{cal.}$	F_{tab}
Head Circum.(x1)	900	0.21	-	-	-	-	-
Body Height (x2)	900	0.19	-	-	-	-	-
Body Weight (x3)	900	0.38	0.31	0.556	0.196	68.39	2.09
Palm Length (x4)	900	0.14	-	-	-	-	-
Feet Length (x5)	900	0.29	-	-	-	-	-
BMI (x6)	900	0.67	-	-	-	-	-

$P \leq 0.05$

Table 2: Regression Analysis Showing the Effect and of Body Parameters on Pupils' Academic Performance in PMSAT.

Sources of variation	N	B	Se	$t_{cal.}$	t_{tab}
H C (x1)	900	- 0.08	0.21	-0.381	1.960
Body Height (x2)	900	0.44	0.19	2.316	1.960
Body Weight (x3)	900	0.2	0.38	0.526	1.960
Palm Length (x4)	900	0.24	0.14	1.714	1.960
Feet Length (x5)	900	-0.25	0.29	-0.862	1.960
BMI	900	-3.04	0.67	-4.540	1.960
Constant	900	45.02	26.46	1.701	1.960

$y = 45.02 - 0.08x_1 + 0.44x_2 + 0.2x_3 + 0.24x_4 - 0.25x_5 - 3.04x_6$

deviation and inferential statistics of correlation and regression analysis tested at 0.05 level of significance.

RESULT

Hypothesis 1

There is no significant relationship between the body parameters and pupils' academic performance in mathematics scholastic aptitude test.

Table 1 revealed that $r_{cal.} = 0.556$ is greater than r_{tab}. = 0.195, at $\alpha = 0.05$ level of significance. Hence, the null hypothesis is rejected; this means that there is a positive significant relationship between body parameters and pupils' academic performance in Mathematics scholastic aptitude test. This showed that pupils' body parameters and academic performance in scholastic aptitude test are positively correlated. The table equally showed that r^2 = 0.31, which means that body parameters could account for 31% of the variability in scholastic aptitude test. This implies that body performance could not account for 69% of the variability in the pupils' academic performance in scholastic aptitude test. This indicates that other variable(s) other than body parameters account for 69% of the variability of pupils' academic performance in Mathematics scholastic aptitude test. The table also indicated that the degree of alienation is 83.07% which means that 83.07% of the variability in academic performance in scholastic aptitude test is strange or alien to body parameters, that is, other variable(s) constituted 83.07% in pupils' academic performance in Mathematics aptitude test. Finally, we observed from the table that F_{cal} = 68.39 is greater than F_{tab} = 2.09, this means that the result of r^2 = 0.31 is not as a result of chance but rather as a result of true relationship between the body parameters and pupils academic performance in scholastic aptitude test.

Hypothesis 2

There is no significant effect of body parameters on pupils' academic performance in mathematics scholastic aptitude test.

From table 2, we can observed that the body parameters of body height, body weight and palm length showed a positive effect on the pupils academic performance in mathematics aptitude test with beta weight of 0.44, 0.2 and 0.24 respectively. While the parameters of head circumference, feet length and BMI showed a negative effect on pupils' academic performance on scholastic aptitude test with beta weight of -0.08, -0.25 and -3.04 respectively. From the regression equation below table 2, indicates that:

(i) For every extra 1cm increase in body height, body weight and palm length, there exist a corresponding increase of 0.44%, 0.2% and 0.24% points in pupils academic performance in Mathematics scholastic aptitude test respectively.

(ii) For every extra 1cm increase in head circumference, feet length and BMI, there exist a corresponding decrease of -0.08%, -0.25% and -3.04% points in pupils' academic performance in scholastic aptitude test respectively.

(iii) For every unit increase in other variable(s) other than body parameters, there is an increase of 45.02% points in pupils' academic performance in scholastic aptitude test.

Furthermore, the table also revealed that the body height is the best predictor of academic performance in scholastic aptitude test and the worst predictor is the BMI.

The t_{cal} for the body height is 2.316 and the absolute value of t_{cal} for BMI is 4.540 which are greater than the t_{tab} = 1.960, this means that the beta weight of 0.44 and 3.04 is the respective exact estimated effect of body heights and BMI on pupils' academic performance in Mathematics scholastic aptitude test and not by chance while the effects indicated by the beta weight of H.C, body weight, palm length, feet length and constant (that is., other variables) are by chance, since their individual t_{cals} is less than the t_{tab} = 1.960.

DISCUSSION

The multiple results of hypothesis 1, revealed that there is a positive significant relationship between body parameters and academic performance in mathematics scholastic aptitude test. This was supported by the findings of a study carried out by Galton (1869) quoted in Miller (1982) were he reported that body measurements is closely related to mental measurements. According to Alonge (2004) all mental test is in reality an aptitude tests. So, the correlation coefficient actually showed the strength of the relationship of body parameters and intelligence quotients of pupils, since performance in aptitude test reflect intelligent quotients. The findings that other variables contributed 83.07% of the variability in academic performance in mathematics scholastic aptitude test is in agreement with the findings of Bandura (1993) who confirmed that other variable such as self-efficacy correlated with academic achievement-related behaviours: including cognitive processing, achievement performance motivations, self-worth and choice of activities. He further asserted that students who are not confident or perceive themselves incapable of learning mathematics may avoid tasks that are seen as challenging or difficult, while those who are highly efficacious will be may willing to face difficult or challenging problems.

Finally, the multiple results of hypothesis 2 revealed that only body height has significant positive effect and body mass index (BMI) has significant negative effect on the pupils' academic performance on mathematics scholastics aptitude test (PMSAT) and there effects are the exact effect they exerted on mathematics scholastic aptitude test and not by chance. This result of body height confirmed the findings of Walk et al, 2000; Kim et al., 2003 and Pearce et al, 2005, who found that the height of children positively correlates with their academic performance and intelligence.

Also the result of BMI on performance in scholastic aptitude test corroborated the findings of Schumacher and Desimone (2005) who stated that weight gain and being overweight and obese each reduces academic performance by a significant amount that is typically statistical. Since BMI is the ratio of weight to height squared in meter, so increase in weight also contributes to increase in BMI. The results of the effects of the independent variables of head circumference, body weight, palm length, feet length and the constant variables (variables other than body parameters) on academic performance in mathematics scholastic aptitude test are as a result of chance rather than the independent variables, since their t-cals are less than their t-tab.

CONCLUSION

Based on the findings of this study, the following conclusions were made:

Body parameters have positive significant relationship with academic performance of pupils in mathematics scholastic aptitude test. That body height has positive significant effect on pupils' academic performance in mathematics scholastic aptitude test. BMI have negative significant effect on pupils' academic performance in mathematics scholastic aptitude test. Head circumference, body weight, palm length and feet length have no significant effect on pupils' academic performance in mathematics achievement test.

Body height is the best predictor of academic performance in mathematics scholastic aptitude test. BMI is the worst predictor of academic performance in mathematics scholastic aptitude test in mathematics.

Recommendation

We recommend that pregnant women should be placed on a health balance diet throughout the period of pregnancy and more especially during the first trimester.

At birth, a child's body height should be measured and if there is an indication of stuntedness in the growth, then growth hormone (GH) could be applied so that the child could catch up growth. The shape and size of the head circumference of the child should be left the way it was at birth, since trying to normalize or reshape the head may likely distort the content of the head.

School management should outlaw the use of abusive appellation with respect to shortness or tallness, or any gesture that could imply that by teachers, pupils or students. School counsellors should introduce programmes to boost the self-confidence, self-image and self-esteem of the pupils regularly. Law makers should fight childhood obesity by making laws that forbid the sale of 'junk foods' in schools; and should regulate the portion size, fat and sugar contents of cookies sale in the state and country. Comments on the cognitive, affective and psychomotor domains of pupils in the assessment report sheet should be mild, encouraging and having no reference or inference to the body parameters of the child.

REFERENCES

Adima EE (1989). Special Education. Ibadan. NPS

to Contents, page 1

Educational Publishers Limited

Ashlesha D, Rolando S, Magna JL (2004). Childhood over weight and Academic Performance: National study of Kindergarteners and first goodness. Retrieved 4/21/2010. Website: http: //www.nature.com.

Brennard H, Fairclough J, Hill V, Nicholson E, Rees E (2005). Child Development. A Comprehensive Text for GCSE. Great Britain, Hodder and Stoughton Educational.

Case A, Paxson C (2008b). Stature and Status: Height, Ability and Labour Market Outcomes. Journal of Political Economy, 166(3): 499-532. Website: http://somesmallsense-blogspot.com2006/08/Princeton-study-correlates-height-and.html

Cawley J, Spiess K (2008). "Obesity and Skill Attainment in Early Childhood", Economics and Human Biology, 6: 388-397.

Conner PD, Sampson PD, Bookstein FL, Barr HM, Streissguth AP (2001). Direct and Indirect Effects of Prenatal Alcohol Damage on Executive Function. Developmental Neuro-Psychology, 18: 331-354.

Encyclopedia(2010). Height and Intelligence. Retrieved 04/28/2010. Website: http://www.statemaster. conferencyclopedia/heigth-and-intelligece.

Fauziah M (2010). Head Growth in Infancy tied to Later Intelligence. Retrieved 5/12/2010. Website:http://www.gateway.n/m.nih/gov/meeting

French SA, Story M, Perry CL (1995). Self-Esteem and Obesity in Children and Adolescents: A literature Review. Obesity Research 3: 479-490

Kim HYP, Frongillo EA, Han SS, Oh SY, Kim WK, Jang YA, Won HS, Lee HS, Kim SH (2003). Academic performance of Korean children is associated with dietary behaviours and physical status. Asia Pac. J. Clinic. Nutrit.

Kolawole EB (2011). Common Entrance Mathematics with Quantitative Aptitude for Primary 5 and Primary 6. Ado-Ekiti. Adedayo Printing Nig. Ltd.

Li X (1995): A study of intelligence and personality in children with sample obesity. International Journal obesity 19: 355-357.

Miller GA (1982): Psychology: The Science of Mental Life. Great Britain. Pengium Books Ltd.

Morgane P, m Austin-La FR, Bronzino J, Tonkiss J, Diaz-Cintra S, Cintra L, Kemper T, Galler JR (1993). Prenatal Malnutrition and Development of The Brain. Neuro Science and Bio-Behavioural Reviews, 17: 91-128.

Mo-Suwan L, Leb els L (1999). School performance and weight status of children and young adoles cents in a transitional metabolism disorder. 23: 272-2 77.

Ogunshola F (2009). Biological Development Parameters of Elementary School Pupils and Academic Performance: Department of Industrial and Technical Education, FUT, Minna.

Owuamanam DO, Owuamanam TO (2004). Fundamentals of Educational Psychology. Lagos. Bolabay Publications.

Sabia JJ (2007): The effect of body weight on adolescent academic performance. Southern Economic Journal. 73 (9) pp 871-900.Retrieved 04/28/2010.Website:http://www.highbeam.com/doc/IGI-163262583.htmI.

Schonfeld AM, Mattson SN, Lang AR, Delis DC, Riley EP (2001). Verbal and Nonverbal Fluency in Children with Heacy Prenatal, Alcohol Exposure. J. stud. Alcoh. 62: 239-246.

Schumacher, E. and Desimone, J. (2005): Academy Health Meeting. Boston Mass,Retrieved5/12/2010.Website:http://www.gateway.n/n.nih.gov/meeting

Stein Z, Susser M, Saenger G, Marolla F (1975). Famine and Human Development: The Dutch Hunger Winter of 1944-1945. New York: Oxford University Press.

Stoch MB, Symthe PM, Moodie AD, Bradshaw D (1982). Psychological outcome and CT findings after growth under nourishment during infancy, A. 20year developmental study, Developmental medicine and child. Neurology, 24: 419-436.

Strauss RS (2000). Childhood Obesity and Self-esteem. Pediatrics, 105-115.

Winick M, Rosco P (1996): The effect if severe early mal-nutrition on celluar growth of human brain. Pediatries research 3: 181-184.

Zhang Z (2010). Childhood Malnutrition could Weaken Old Age Brain Function. The Punch p. 70.

to Contents, page 1

PRIME Journals
Prime Research on Education (PRE)
ISSN: 2251-1253
Vol. 2(1), pp. 147-154, February 3rd, 2012
www. primejournal.org/PRE

Full Length Research

Adopted business model for mobile learning

Behrouz Zarei[1] and Maryam Safdari[2]*

1Faculty of Entrepreneurship, University of Tehran, Tehran, Iran.
2Lulea University of Technology, Sweden.

Accepted 7th January, 2012

Relying on the use of mobile device which is increasingly popular worldwide, mobile learning in fact extends the reach of education to all social-economic levels independent of location and time, indicating a new opportunity for education industry development. Nonetheless, there is still a lack of a comprehensive understanding of the influencing factors on the adoption of mobile learning. In this study, an adoption mobile learning model was built which is based on the business model ontology, in which, the exploratory qualitative study used case study method. Items were identified from the data and defined according to their general properties and dimensions. We recognize the main elements of mobile learning business model. According to mobile learning infrastructure and its new approaches in developing countries, sub-elements of current business model modules are varied in comparison with in developed countries. Therefore, it requires many relevant activities to support adopted sub-elements that should be attained target customers, maintained and supported them. This model hopefully provides a framework for future research and business.

Keywords: Mobile learning, business model.

INTRODUCTION

As an emerging paradigm in a long tradition of technology-mediated learning, mobile learning is defined as the acquisition of any knowledge and skill through the use of mobile technology, anywhere, anytime that results in an alteration in behavior (Geddes 2004). In recent years, there has been great interest in the potential for mobile learning. Some providers of m-learning components achieve profits by offering m-learning products. They are likely to become key players in the field of m-learning. Others developing with the same or even more efforts do not succeed in supplying m-learning products.

Currently, mobile learning is emerging as a promising market for education industry. On one hand, from a technology perspective the tipping point for mobile learning is coming closer as technology improves and standards emerge (Quinn 2008).

Mobile learning posits unprecedented opportunities for both education institutions and governments as well. In the context of education institutions, many higher education managers have seen mobile learning as a way of extending the reach and hence increasing revenues (Murphy 2006). Services acceptable and to be Whilst there is a growing interest from both academic and business communities, the issues regarding how to promote learner's adoption of mobile learning seem to be largely unsolved, and thereby posit to be a challenge for services providers. For instance, according to Corbeil and Valdes-Corbeil (2007), students access to various mobile devices can not ensure that their ability to achieve educational goals it seems an urgent need to recognize effective factors of user intention treatments to preserve and create user fees and developing cost (Yong Liu et al, 2010).

The business model is part of a comprehensive m-learning strategy. A business model determines strategic functions with considering economical aspects, pedagogical and technical dimensions for m-learning providers. (M, I.S., M. Hosany and R. Gianeshwar, 2006).

"Mobile Learning has a strong foundational base when it comes to how the student will learn when there is an interaction of the learning material, technological platform, and the wireless network. Just like wireless technology which was built on numerous technological advances, M-Learning also is a combination or a hybrid of more than one system. In this case, M-Learning relies

to Contents, page 1

on pedagogical theories and strategies of the Behaviorist, the Cognitive, and the Constructivist Learning groups". (Mobile Learning Groups, 2004).

Only if all parties (mobile customers, network operators and service providers) can increase their profit significantly, the whole market potential might be realized. Therefore existing mobile business models have to be reconsidered with regard to, the following elements, Timmers (1998):

- Value proposition,
- Key customers
- Core activities
- Business partners
- Revenue flows

This element is a guideline for orienting research. As a first step towards developing a new business model, one can start studying a current mobile business model and identify its shortcomings.

There is no single look at the mobile learning. Actually a deep gap can be seen between researchers and policy makers in the field. In fact, effective and efficient framework and the relationship between them have not been defined and established. In this context, this study aims to develop a good and integrated business model that establishes the required relations to be used in the operation area.

Considering to recent technology advances particularly in mobile, redefining learning concept and framework becomes necessary to recognize the role of mobility and communication in the learning process, and deformation in the digital networks to protect virtual social networks and removing their obstacles (Martinovic et al, 2010).

All PC-based learning content cannot convert into a mobile format, but to consider how the mobile devices can be used to make efficient learning strategy through mobile format and devices. Reasoning is brought by Advanced Distributed Learning (ADL), which is help to organization to enable the interoperability, accessibility and reusability of Web-based learning content (Bayoumi, 2007).

In addition, the possibility of mobile learning can be explained from market trends. The manufacturers are providing various mobile devices now. A significant number of people are already using these mobile devices to access internet and use their documents, and so on.

Business model

A company's business model is simplified representing its business logic. This describes how a company provides its customer need from initial to the end point through potential and applicable resources, activities, business partners and makes money (*Kemp, 2009*). The business model is usually different from the business process model and the organization model. Business models can be described in a more or less formal way but A Business Process Model (BPM) is commonly a diagram representing a sequence of activities from end to end.

"Conceptualizations of business models attempts to formalize informal description building blocks and their relationships "(A. Afuah and C. Tucci. 2003). In this regard we use a simple approach that consists of nine basic business model building blocks that allow us describe and draw all the aspects of a business model in a simple way.

First, core capabilities are the necessary capabilities and competencies to run a business model in the company. Second, trade unions, which supplements other aspects of business model as a network partner. Third, the logic makes for mutually beneficial business and its customers are value configuration.

Once a business provides the products or services then value proposition is appeared. Quoting Osterwalder (2004), a value proposition "is an overall view of products and services that together represent value for a specific customer segment. It describes the way a firm differentiates itself from its competitors and is the reason why customers buy from a certain firm and not from another."

Distribution channels offer by the company to deliver products and services to its business target audience for products and services. Along marketing, distribution strategy of business and customer relationship can be created the link between company and its different customer segments. The process of managing customer relationships is referred to as customer relationship management.

Last section is business finance, including cost structure, the consequences of monetary business model and means of employment income and the way a company makes money through a variety of revenue flows a company's income.

METHODOLOGY

Ineffective and costly traditional learning makes clear the needs of E-learning and subsequently mobile learning. Some causes of inefficient learning system can be mentioned as follows;

- separate learning environment from actual environment,
- disintegrated business processes
- Impossibility to receive necessary learning sometimes as a result of the appearance of problems
- Requirement of immediate learning and knowledge.

There is no single look at the mobile learning. Actually a deep gap can be seen between researchers and policy makers in this field. In fact, effective and efficient framework and the relationship between them has not been defined and established. In this context, this study aims to develop a good and integrated business model that can reduce the above weaknesses and establish the required relations to be used in the operation area.

Despite existing shortages, lack of an integrated mobile learning business model as an instrument is evident for transferring many potential opportunities to actual ones in this area. This research project intends to develop the best possible M-learning business model for Iran. To deliver on the potential of mobile learning, it is needed to consider these main questions:

- Which types of business model are best suited to mobile learning?
- What components can be in this business model?

This research is empowered with two points of view realism and subjectivism. Once we follow to realize a business model concept it is looked within subjectivism point of view and when we want to specify our business model for mobile learning area we wear realism glasses.

Moreover, this research as an interpretive study is established on Hermeneutics and emerging nature. Generally interpretive considerations attempt to percept through phenomena sense of individuals. Interpretive research does not define dependent and independent variables in advance. As a matter of fact this research implies some parts of this philosophy and attempts to study on reality to enable provide business model expansion.

In case of similar problems to this subject, the research method will rely on former specified and step by step methods, but for the essence of new and complicated problems, the previous methods must be reviewed even to make a new research methodology. This subject is consisted of management, electronic commerce and their practical subjects. To sum up intellectual questions and cognitive disturbance, it was indicated that study on business models area can be respondent to researchers internal requirements and has adequate strength in researchers provocation. So for this research follow process has been lunched:

- Select appropriate subject proportionate to our internal requirement
- Evaluation and design of research methodology and reliable process that facilitate researcher attain her goals by a certain method
- Initiation of business model perception and explaining it in mobile learning industry
- Attempts to necessities recognition, applications mobile learning and business model specialty and reach to customized model
- Explanation of extract model and its relations
- Make reliance of results and publish them

This research direction also in search of clear space study can be called an exploratory research. Especially during the exploration part of the transparency in the existing business models focuses on mobile learning. Research strategy of this study is case study that it has been selected two cases in the area of virtual/mobile learning industry which is relatively complete range of material discussions on issues.

According to the few Iranian active companies in mobile learning, these two case studies were closer to the subject and they could cover relatively complete spectrum of subject materials in our discussion then open ended interviews were held for them. Before the data collection different business model references has been studied and one near business model (Business Model Ontology, Osterwalder 2004) is selected as our predicted model. The main aim was approached from the studied phenomenon to the predicted model. With the help of Grounded Theory idea, the interview concepts were open coded, reviewed, compared and launched.

Placements of launched concepts into predicted business model then subsequently the relevant relationships were determined between categories in the business model.

Business model considerations for M-learning

Unlike the Internet business model classification, mobile business is classified service model with mobile features and also Internet business features.

The mobile business model classification scheme has a couple of differences from the previous research. First, the classification is based on a business model perspective: previous mobile business classifications have been focused on service categorization, not representing the way business is run on the mobile channel. Service categorization has its own limitations to express mobile business. The new mobile classification scheme is based on a business model perspective. In fact, mobile learning offers another way to deliver content and to embed learning into daily life. The learning materials need to be developed in small, consumable bytes of format, which can be delivered through wireless network.

Customer Segments

Corporate/Business drivers

The corporate mobile learning grows louder with each day. Organizations no doubt recognize that mobile technology for learning has merit. In the corporate arena, many recognize the need for just-in-time training and performance support and are beginning to explore mobile options. In the medical, sales, and service areas, mobile is being used most widely. Specific projects related to m-Learning from the corporate side have involved organization.

Education

Mobile learning products that are fun, interactive, and instructionally sound is suited for this segment. The

to Contents, page 1

instructional designers, content developers, animators, graphic designers, and child psychologists has designed interactive learning products for children of diverse age groups, from kindergarten to senior high school.

Most of Iranian mobile learners are students and literate level otherwise there are in the universities or those of students who prepare to enter universities through a national exam. Thus mobile learning products and services for this segment can add value to the curriculum, allow students to apply their learning in real-world situations, and cater for different learning styles. The solutions address these issues; they are also flexible enough to incorporate the differences between cultures, curricula, and instructional methodologies.

For mobile learning to enter into the mainstream the following are necessary:

- Guides for the development of mobile learning need to be produced for mainstream institutions
- Mobile messaging needs to be established as a standard administrative structure in education and training institutions; short courses, course summaries, examination preparation, needs to be developed for PDAs, smart phones and mobile phones.
- Full modules need to be developed for PDAs and smart phones; these modules need to be offered to fee-paying students, as part of the institution's offering, with full evaluation and with normal awards and accreditation.

Government

The highly skilled in creating attractive and interactive yet accessible products that have delivered products. It can be recognized that accessibility is both urgent and imperative, especially in the government space, and it is fully geared to meet the associated developmental challenges. The capabilities also include the conversion of mobile content into accessible material.

The solution is for the requirements of clients in the government, and local and defense areas. The Subject Matter Expertise and content developmental skills add up to high quality learning content, both Web- and CD-based, in a wide variety of areas.

Value proposition

M-Learning can add value to the experiences of many learners but it may not add value equally to all. This does not justify denying those who can benefit but care must be taken to ensure others are not substantially disadvantaged. The section below illustrates a number of ways m-Learning can benefit all learners and in the process bring particular advantages to different types of disabled learner. To a disabled learner the added value of m-Learning is three fold:

- Any assistive technology benefit is more portable so the support available to the learner is available more places and more times.
- Mobile technologies are generally cheaper than PCs and laptops so more likely to be affordable.
- Mobile technologies are private and personal in use and have none of the student self-image problems that may be associated with traditional assistive technologies.

New services and new tools: Iran service providers offers an interesting products that emerged in 2008 was the availability of several specialized tools designed to create very specific types of applications such as quizzes, flash cards, tourist guides, language learning content, and performance support.

Standard Products and Services: These service providers state that they produce contents and develop their software themselves basis to their customer request specification so their products is specific to Iranian customers.

Reducing Cost: During the interview with experts is earmarked that presenting new products and services to mobile learning customers can reduce significantly time and cost because the data is transferred with fast rate and subsequently there is no need to present some learning material that are mobile potential in traditional learning. Thus it can reduce cost and create value.

Monopoly License and Software: The mobile service providers in Iran produce mobile learning products in their own License and also they develop necessary and related software according to their customer requirements.

Content Management: From these cases is cleared mobile learning providers produce and manage some their mobile learning contents. Most of the content was ported from other mobile formats, yet several of the largest educational publishers entered the market for the first time through retailing.

Distribution channel

Content delivery to mobile devices may well have a useful place in m-learning; however, there is an imperative to move from a view of e- and m-learning as solely delivery mechanisms for content. Packaged Mobile Learning content will sell through below primary distribution channels:

Retail: Packaged content products will available on dozens of retail sites. The retailers reached a relatively small number of customers but they can drive the market with online retail content stores that reach a vast number of customers. The launch of these stores has given developers a direct to consumer channel.

SMS: All higher and further education institutions have a frequent need to provide information to their students about timetable changes, assessment deadlines, feedback from tutors and other urgent administrative details.

E-mail: Some products and learning materials delivers to

Iranian mobile learners through Email.

Mobile Internet: Iranian service providers extend their distribution channel through shift technology to 3G it is optimum method that users can access to the internet through a mobile devices and capture their materials and learning contents.

E-files: this channel is as a usual way to deliver packaged mobile learning in Iran.

High Rich Media: Mobile learning providers produce video conference infrastructure and context for some universities and governmental organizations such as Lorestan Universities and Hygiene and Medical Education Ministry.

Brand: There are a few brands in mobile learning field in Iran that have influence significantly in mobile learning content

Relationship management

The mobile operators need a scientific method to guide the customer acquisition. The method must be able to identify the target customers precisely and analyze the customers' actual needs. And the mobile operators also need an information system to support the whole process of customer acquisition, which can improve the efficiency and effectiveness of acquisition.

For the Iranian customers, the operators can find some promotional ways to upgrade their consumption levels, which could bring more values to the organization. This depends on understanding customers' behavior, the abilities of innovation of new services, and marketing capabilities through data mining technologies that it supports customers from start point through their personal relationships.

The type of relationships can shape and organize to lunch customer acquisition and retention. Imagine exposing a highly tailored series of M-learning modules to specific end-users. Companies already sponsor luncheons, conferences and workshops to expose their products to physicians. So why not use M-learning to do this to reinforce these efforts? By providing M-learning to the customers, it can be increased customer loyalty and potentially grow the revenue, through customer acquisition.

Revenue streams

It is essential to identify potential revenue streams to prepare the institution for 21st century learning. The possibilities will vary widely from institution to institution, for example. realising capital locked in city centre sites, acquiring funding from urban or rural regeneration schemes, using development funding, or tapping into new markets. Consider also:

- Ways in which mobile and wireless technologies could improve the financial standing of the institution by raising its status in the community, improving efficiency and increasing retention.
- How increased use of personal mobile devices could offset the institution's costs
- *License Fees*; the Iranian M-learning providers sell some parts of their product through selling license in which most of them are done by retail products
- Companies can send their *advertisements* to students on mobile devices along with the study materials which is send to them. They can give exact details of their product and their availability. As this advertisement come along with the study material they would surely have a look at that.
- Iranian providers define four types of membership to receive learning packages and alerts on time and presenting some promotions during membership on mobile devices
- Individual membership (academic institutions such as Gaj and Ghalamchi)
- Student in full-time membership (schools and universities)
- Education Provider (medicine school)
- Commercial Organization (language learning providers)

Key resources

Mainly the impact of technology, environment, market and regulation, network, wireless network, content and device issues have a direct influence as our key resources in business model.

Environmental drivers means technology, regulations and market have a direct influence on our key resources in our business model.

M-learning providers in Iran state that they have pressed by the current regulations in this industry which they issued by governmental sections. They establish strict internal regulations for them that are effected by influential relationships by third parties. According to existing shortages in copy right in Iran these providers are not in safe side to produce packaged mobile learning in their own license.

Business and business model environment in practice are not static, m-learning business like the other electronic and telecommunication businesses put into practice in a rapid change environment and market, this business changes may be ascribed to internal forces, they are caused by external events as well.

With reference to the technological aspect, handheld devices, content and wireless access technologies are the most important drivers forward the change in m-learning business models (Sanz-Velasco, S.A., 2007). Five broad categories of technology were identified, within which specific technologies would need to be selected, these were: transport, platform, delivery and media technologies plus development languages.

The mobile learning network is a unique collaborative approach to encouraging, supporting, expanding and promoting mobile learning, primarily in the English further

to Contents, page 1

education sector, via supported shared cost mobile learning projects. Collaboration at national level involves participating institutions and sharing the cost of projects introducing or expanding mobile learning providing a support and evaluation program.

Business cannot be devised without an understanding of the needs of potential customers of valued added m-learning (Parsons et al, 2006). The scope of mobile learning which help the supplier understand the *market* and focus on to gain the large quota of market.

M-learning regulation and standards is a companion document to the m-learning education and training system and drive the way m-learning can be used in teaching and learning. The m-learning regulations cover a general strategy and road map for m-learning and its standards cover topics.

Content providers offer prearranged m-learning content. Content is either standardized, or personalized or individualized. This content could be data and information products.

Probably the first device that comes to mind when mobile hardware is discussed is the PDA. These devices offer many of the features of a full-size laptop computer but in a package that fits in a pocket

Key activities

Software development

Iranian software developer is presented the methods by which a practitioner, with limited technical knowledge and time, might produce learning objects for mobile phones by adapting existing materials that is selected three:

- A software which are used to code web pages
- A software which is freely available for educational use
- A mobile movie format called Third Generation Partnership Project

Key partners

Iranian partners in mobile learning industry include;

- Network Operators: Hamrahe Aval, Irancell
- Holding Companies and Investors: Governmental organizations such as Radio and Television Organisation
- Software Providers: Soroush Hamrah, Souroush Data
- Third Party Partners; Universities Government

Cost structure

New phones are capable of exchanging voices, text, pictures and video. In addition, the wireless network provides high-speed connections with low costs to mobile subscribers.

Outsourcing: Information Technology (IT) governance is concerned with decision rights, resource mobilisation and strategic alignment to achieve more benefits from information systems. These mentioned providers outsourced the M- learning development it would be a quicker way from the idea, to having it ready for testing, but to integrate it with our servers and other services it could turn out to be problematic. There is needed to be able to control when and to who want to send m-learning products.

Customer acquisition cost: An important business metric, customer acquisition cost should be considered along with other data, especially the value of the customer to the company and the resulting return on investment (ROI) of acquisition. The calculation of customer valuation helps a mobile business decide how much of its resources can be profitably spent on a particular customer. One of these cases shows that has much customer acquisition cost in comparison with other case because the second one is the subsequent companies of Radio and Television Organisation in Iran and this organization optimize this cost for them through their different advertisement methods as well they have their own customers that can expand them to mobile learning customers as a pioneer in content providing in Iran.

Personnel Cost: Their personnel cost focus on salaries, employee's insurance cost, training and providing leisure activities for them (Figure 1).

CONCLUSION

Recently the use of the Internet business has been very prevalent. There are world examples of successful managers who are using new models and have been able to succeed despite intense competition in the labor market, creating a safe place for their company among the competitors there.

This study explores to develop mobile learning business model that is able to present mobile learning Profile Products for Iranian customers of this industry. In addition, the manner of model presentation, internal relations management with market and products delivery conditions has been created to achieve stable market share and increase revenue flow.

This business model can be the ring strategy and business processes together and also can offer an example of mobile model's elements in connection to each other which are:

1 - Product innovation
2 - Type of relationship with customers
3 - Management Structure
4 - Methods of financing

We think when talk about the business model, you should look up to the elements that can be answered the above four-question, the answer is given or not? In addition, each of these elements, have some subset of them should be considered.

Most of mobile business models are just the concepts without more discussions, the least understanding and consensus for them and this study attempts to cover

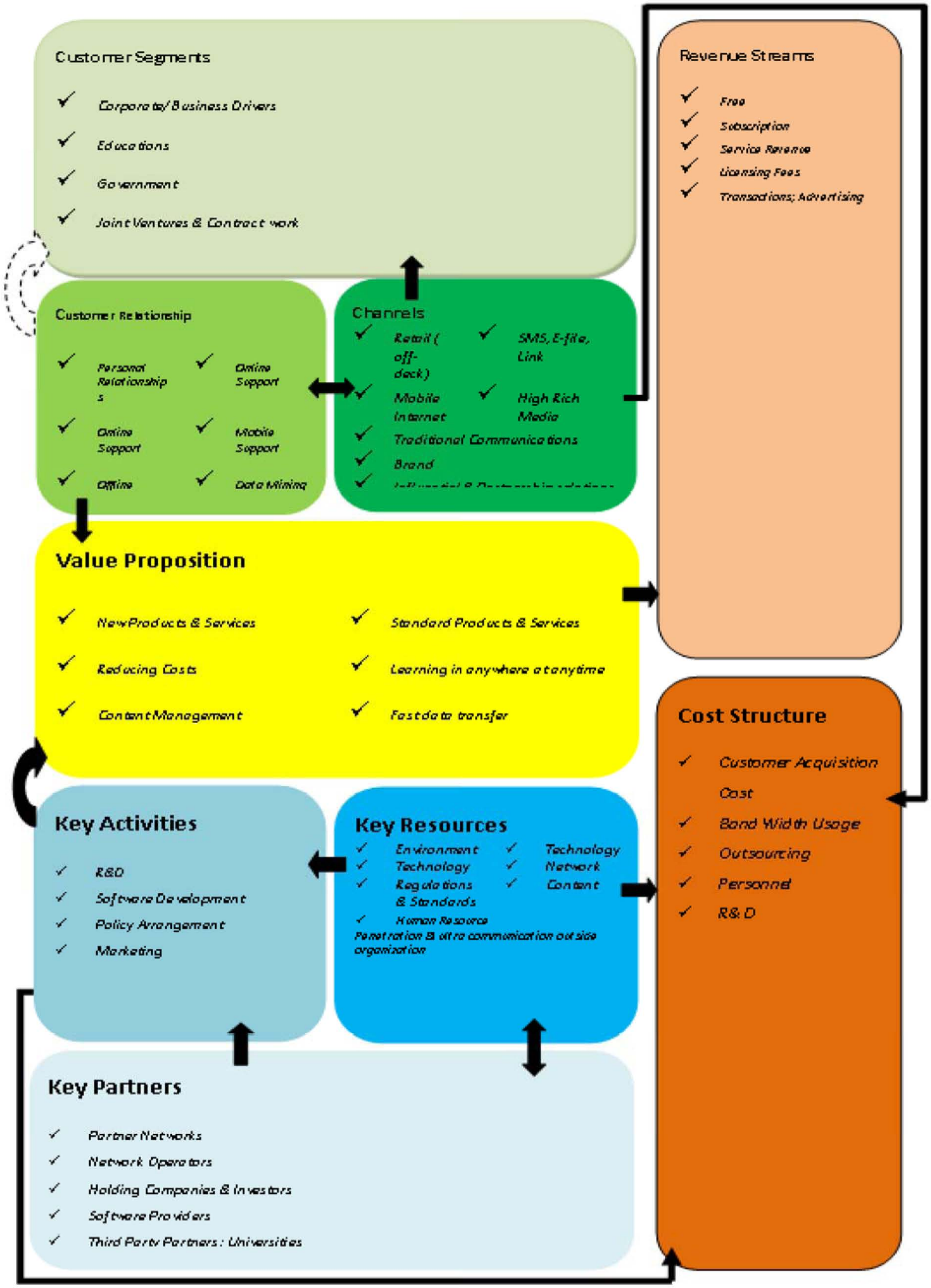


Figure 1: Please provide legend

to Contents, page 1

these shortages. By this we can state this model can create new values through any direct and indirect money exchange by mobile networks and simple methods in mobile learning business to maintain stability of in market and revenue flow.

So that should be mentioned that this study wants to proceed electronic specially mobile network as a family of standards that is allowed to participants in the mobile learning business to find out truth of each other and makes communication connection based on unique basis like special framework that enables us to answer mobile learning business questions and ambiguities.

During this study we faced to some problems that it is caused we found the classic and traditional business models structures have many contradictions with today's conditions. We think it is a rational way that managers must use purposeful and effective management techniques to become success in the future.

Generally, along this study researcher found some business models due to poor value objectives and value purposes defeat. This may happen when companies cannot understand the place of value in value chain.

This research experiences marks disability in balance with what should be done through internal capabilities, and what should be outsourced is the main reason of rejection a business model. Specific capabilities for process strategy may be obtained from existing sources may be the formation of a big mistake because useful. Moreover a next issue in business models is mismanagement and tact in the values network. Value network management can be an important factor to become successful.

REFERENCES

Afuah A, Tucci C (2003). Internet Business Models and Strategies, Text and Cases. 2nd Edn., McGraw-Hill, New York, ISBN: 0072511664.

Bayoumi F (2007). Guidelines for Developing Adaptive Mobile Learning. Conference IMCL2007.

Corbeil JR, Valdes-Corbeil ME (2007). Are you ready for mobile learning?, Educause Quarterly, Vol. 30, No 2pp. 51-58.

Geddes SJ (2004). Mobile learning in the 21st century: benefit for learners. Knowledge Tree e-journal: An ejournal of flexible learning in VET, 30(3): 214-228.

Kemp N (2009). http://coherencymanagement.org/ coherency-management/cases/winnipeg-fleet-management.

Martinovic D, Pugh G, Magliaro J (2010). Pedagogy for Mobile ICT Learning Using Video-Conferencing Technology. Interdisciplinary Journal of Information, Knowledge, and Management, 5: 375-95.

M IS, Hosany M, Gianeshwar R (2006). A Mobile-E-learning (M-E) adaptive architecture to support flexible learning. Malaysian Online J. Instruct. Technol., 3: 19-28.

Mobile Learning Group, 2004 MDDE 663; http://adlib.athabascau.ca/resources.htm#ped.

Murphy A (2006). Mobile learning in a global context: a training analysis. Fifth International Conference on Networking and the International Conference on Systems (ICNICONSMCL'06). Morne, Mauritius, pp. 219-219.

Quinn C (2008). What's the (m) point? http://blog.learnlets.com/wp/?p=287, retrieved on: March 10, 2008.

Osterwalder (2004). The Business Model Ontology - A Proposition in a Design Science Approach.

Sanz-Velasco SA (2007). Technology and business model learning leading to growth: Start up ventures in mobile internet. Int. J. Techno-entrepreneurship.

Timmers P (1998). E-Business models for electronic markets, Electronic Markets, 8(2): 3-8.

Yong Liu, Shengnan Han, Hongxiu Li (2010). "Understanding the factors driving m-learning adoption: a literature review", Campus-Wide Information Systems, 27(4): 210-226.

PRIME Journals
Prime Journal of Business Administration and Management (PJBAM)
ISSN: 2251-1261
Vol. 1(12), pp. 375-381, November 5th, 2011
www.primejournal.org/BAM
to Contents, page 1

Review

The study on the export diversification and constraint in export diversification which volatile the export earning: A case study of Pakistan

[1]Mehwish Bhutto and [2]Sobho Khan Jamali

[1]Lecturer, Department of Economics, University of Sindh, Jamshoro
[2]Assistant Professor, Department of Economics, University of Sindh, Jamshoro

Accepted 23rd November, 2011

It is frequently suggested that export diversification contributes to an acceleration of growth in developing countries. Both horizontal and vertical export diversification are positively correlated with economic growth. Diversification is directly related to the structure of the economy and how it changes as development takes place. We must consider the relationship between economic structure, diversification, and export earnings. Yet there have been remarkably few empirical investigations about export diversification. The current paper is aimed to evaluate the export diversification (both with respect to markets and products) during the FY02-07. Moreover, an attempt is also made to highlight the major constraints in export diversification There are many methods which can be used to analyze the export diversification. In this report my study is based on data on exports to different countries. I have used Lorenz curve and tables as a tool to analyze the export diversification. In this study the purpose of using Lorenz curve is to analyze that whether our exports are concentrated in few regions or these are diversified.

Key words: Export diversification, Lorenz curve, economic structure, Pakistan.

INTRODUCTION

Pakistan is facing trade deficit since the day of independence. The continuously expanding trade deficit touched US$ 13 billion during FY07. The higher import growth as compared to that export has been the dominant factors behind this persistent widening of trade gap. The concentration of exports in few products and markets has been one the major underlying weakness of Pakistan's export which is harmfully affecting our external balances.

The export data suggests that nearly 60 per cent of our exports are cotton based and around 75 per cent are based on five products cotton, leather, rice, textiles and sports goods. In geographical terms, our exports are concentrated in few markets including USA, EU, U.A.E, Japan, Hong Kong and Saudi Arabia. Only USA takes around 24 per cent of our exports. Such a situation shows a serious vulnerability of our economy to a terrible event such as cotton crop failure.

The diversification of exports thus should be the primary objective of the government. The instability in export earnings is very crucial to policy makers of any country. They tend to equate a narrow export commodity base with the instability and declines in earnings, and to propose export diversification as a useful cure.

According to the study conducted by Rizwan Ali; export diversification, by definition, involves changing the composition of a country's exports mix. However, the impact of alternative export mixes on export earnings performance is unclear. In addition, diversification is directly related to the structure of the economy and how it changes as development takes place. We must consider the relationship between economic structure, diversification, and export earnings.

The current report is aimed to evaluate the export diversification (both with respect to markets and products) during the FY02-07. Moreover, an attempt is also made to highlight the major constraints in export diversification

The report is organized as follows, part consists of introduction, followed by methodology, and literature review, in next part concept of diversification is discussed. Followed by data analysis and constraints in exports diversification. The final section contains results and policy recommendation.

Literature review

There are many studies which are conducted regarding exports diversification some of them are given as follow:

Dierk herzer (2004) "Export Diversification, Externalities and Growth"

This paper attempts to examine the hypothesis that export diversification is linked to economic growth via externalities of learning-by-doing and learning-by-exporting fostered by competition in world markets. The diversification-led growth hypothesis is tested by estimating an augmented Cobb-Douglas production function on the basis of annual time series data from Chile. Based on the theory of co integration three types of statistical methodologies are used: the Johansen trace-test, a multivariate error-correction model and the dynamic OLS procedure. Given Chile\'s dramatic changes in economic policy, time series techniques considering structural breaks are applied. The estimation results suggest that export diversification plays an important role in economic growth.

Federico Bonaglia (2003) Export Diversification in Low-Income Countries: An International Challenge after Doha; this paper discusses major policy issues related to commodity dependence and export diversification in low-income countries. Contrary to some widely-held view, it argues that natural resources are not necessarily a 'curse' that condemns low-income countries to underdevelopment but can provide a basis for sustained export-led growth. Natural resource-based sectors have potential for export diversification. The OECD 'mirror' trade data indeed suggest that many different routes to diversification exist, including resource-based manufacturing and processing of primary products. However, these opportunities are far from being exploited in many low-income countries. This is because export diversification is typically a slow process, and this process needs to be sustained by an appropriate and coherent strategy, characterized by a combination of vision, co-ordination and management of conflicting interests.

Akram Khatoon (December 2001) "a strategic approach to exports": in this study it is suggested that apart from protecting existing volume of exports of cotton , rice, leather, sports and so on we should diversify our exports. Major reliance on few items need to be reviewed. According to this study exports diversification would not only boost our exports but also would give impetus to small and medium size industries, as a result employment in the country will increase.

Dr. Ashfaque Hassan Khan (2003-04) "Pakistan exports: what needs to be done?" this paper suggests that for export diversification we should formulate a policy frame work in which the government should design and Support a steady macroeconomic policy framework consistent with export promotion Strategies. While export diversification programs should be implemented primarily by the private sector the role of the government, in this context, should be to check Distortion and create an environment which promotes diversification.

Iran Shahzad (2003) "Diversifying export markets": this study analyzes that exports diversification is necessary to capture a greater share in the global markets for our products. This struggle becomes even more important in order to survive in the "WTO" regime. Export diversification is not the responsibility of exporters only. The government should also play a role in this regard.

Mubarak Zeb Khan (2007) "Causes for slow performance in export identified" according to this study the main reason for sluggish performance of Pakistan's exports is the lack of exports diversification for products and markets. In this study it is suggested that, we should form a policy which focuses on developing strategies for diversification and increase the exports competitiveness. We should devise a policy for increasing the production of various products, not only the production of few commodities (textile rice, and so on).

Zafar-ul-Hassan Almas (Nov 2003) "Export diversification and sustainable growth";

In this paper it is suggested that export diversification is an engine of growth for a developing country. It protects a country from uncertain changes in foreign trade, it is necessary to stabilize domestic incomes and employment. it is opportunity for Pakistan to concentrate on export diversification in both products and direction of trade. Diversification of exports is helpful for long term sustainable growth. Pakistan needs policy fully compatible with the challenges of globalization.

METHODOLOGY

There are many methods which can be used to analyze the export diversification. In this report my study is based on data on exports to different countries. I have used Lorenz curve and tables as a tool to analyze the export diversification between FY02 and FY07 by using Microsoft excel. Basically Lorenz curve is used to analyze the income distribution among the masses. It tells us that whether income is equally distributed or it is concentrated in to few hands. In this study the purpose of using Lorenz curve is to analyze that whether our exports are concentrated in few regions or these are diversified.

CONCEPTION

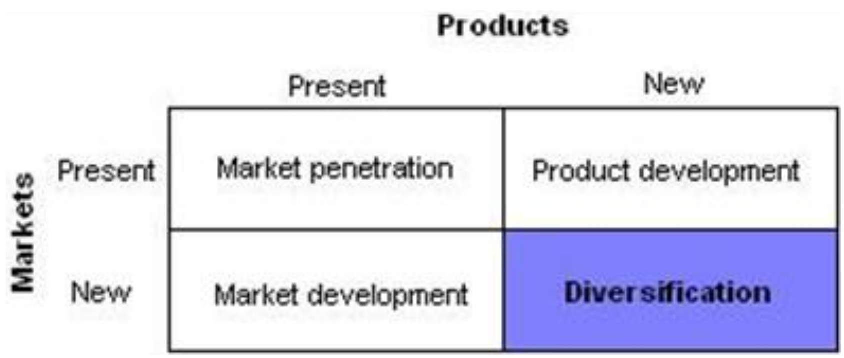
Diversification

It is a risk management technique that mixes a wide variety of investments within a portfolio.

Diversification is part of the four main marketing strategies defined by the soff matrix (Figure 1). According to this matrix the firs three strategies are usually pursued with the same, technical, financial, and merchandising resources used for the original product line, where as diversification usually requires a company to acquire new skills and new facilities. Therefore, diversification is meant to be the riskiest of the four strategies to continue for a firm.

to Contents, page 1

Figure 1. Four main marketing strategies defined by the an soff matrix



Export diversification
It means changing the composition of products which are exported by one country to another country, export diversification can be either by products or either by countries.

Types of export diversification
There are mainly two types of export diversification;
1) Horizontal diversification.
2) Vertical diversification.

Horizontal diversification
Horizontal diversification means adjustments in the export mix in order to counter international prices or export quantity instability. It means to increase the number of commodities.

Vertical diversification
Vertical diversification involves creating additional uses for existing and new commodities through value addition such as processing and marketing. Simply it means movement from low to high value addition.

Advantages of export diversification
Export diversification has many advantages; some of them are given as follow.

Protection against price fluctuations
Export diversification is a safeguard against price fluctuations. If the exports of a country are concentrated into few products, than any reduction in the price of these commodities will adversely affect the monetary value of exports. However if a country exports a large number of products non of which has a major share in its total export earnings, than any reduction in the prices of few export items will not adversely affect the monetary value of exports.

Protection against demand fluctuations
Export diversification protects a country against demand fluctuations in the international markets. If a country exports few products than any reduction in the demand for these products will substantially affect the volume of exports.
However if a country exports a large number of products then the repercussions of fluctuations in international market demand can be minimized.

Check against market saturation
Diversification is a check against market saturation. A country having few export markets and products may face a reduction or stagnation in its market share if the markets become saturated.

Strategy against product life cycle
There are four stages in product life cycle introduction, growth, maturity, and decline. Profits and sales are highest when a product is at the stage of growth. One advantage of export diversification is that the same product may be at different stages in different markets, thus a fall in profits and sales in a mature market can be counterbalanced by increase in sales and profits in growth market.

Improvements in (BOP) position
Export diversification is instrumental for improving the position of balance of payments. As Pakistan's (BOP) position is not good, because our exports are less than imports. In order to narrow down this gape export diversification is necessary. Because export diversification leads to increase the exports earning of a country, so as a result of diversification our (BOP) position will improve.

Data analysis
Lorenz curve
A cumulative frequency curve showing the distribution of a variable such as population against an independent variable such as income or area settled. If the distribution of the dependent variable is equal, the plot will show as a straight, 45° line. Unequal distributions will yield a curve. The gap between this curve and the 45° line is the inequality gap. Such a gap exists everywhere, although the degree of inequality varies. In the graph given below

Figure 2: Lorenz curve showing the percentage between income and households

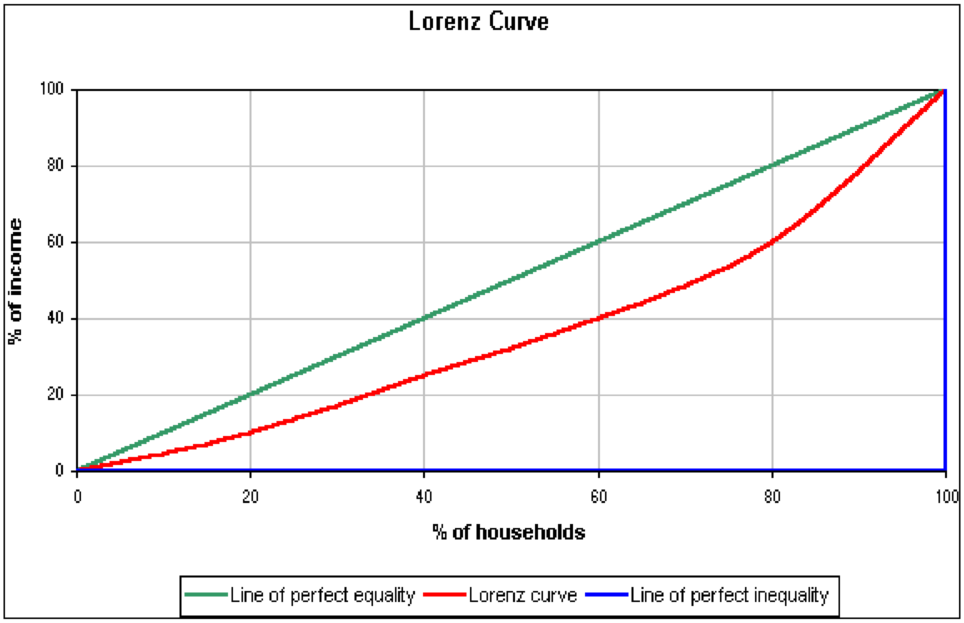
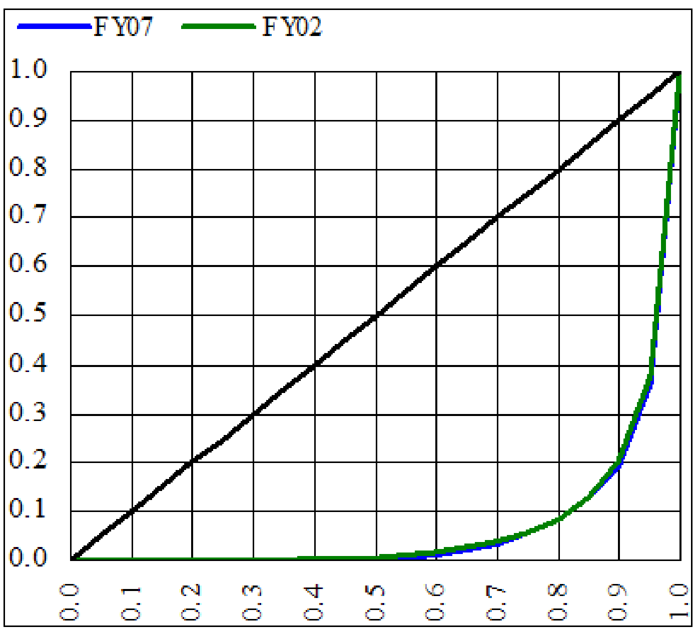


Figure 3: Lorenz curve for export markets



the percentage of households is plotted on horizontal axis, and percentage of income is plotted on the vertical axis. Figure 2.

In the subsequent figure Lorenz curves for export markets are drawn for the FY02 and FY07. Blue Lorenz curve is drawn for the fiscal year 2007, while green Lorenz curve is drawn for the fiscal year 2002. In the following graph there is a 45° line which is indicating the situation of perfect diversification. In the subsequent graph it is obvious that our exports are highly concentrated into the few regions, because there is a great deviation between diagonal and the Lorenz curves. We can also observe that as the Lorenz curves for the FY02 and FY07 are overlapping there is no relative diversification between FY02 and FY07. It depicts that the current situation is not encouraging we are stagnant at the position where we were three to four years back.
We can also analyze with the help of ginni coeffient that whether our exports are concentrated or diversified, as we know that ginni coefficient is the ratio of area between Lorenz curve and the diagonal to the area below the diagonal. In the above graph we can analyze that value of ginni coefficient is higher which shows concentration not the diversification of exports.

Interpretation

As it is clear in the figure 1 that there is no exports diversification. Our exports are highly concentrated into few markets. The results depicted in figure 1 can be supported by the data shown in table 1. In the given table we can observe that in Fy03 that almost 84% of our textile exports are going to top ten countries. While only 16% of our textile exports are going to the rest of the world. Now if we compare it with FY06 than we come to know that situation is no more different, again almost 85% of our textile exports are going to the top ten countries while only 15% of textile exports are going to other countries.
Now if we consider our total exports than again situation is almost akin to the previous one. Because if we observe in the table we can realize that in FY03 approximately 80% of our total exports are concentrated in the top ten

to Contents, page 1

Table 1: Pakistan's Total Exports and Textile exports (Including raw cotton) Market (Million US$)

Countries	FY06		FY05		FY04		FY03	
	% Share in		% Share in		% Share in		% Share in	
	Total	Textile	Total	Textile	Total	Textile	Total	Textile
	100	100	100	100	100	100	100	100
EU-25	25.9	29.3	34.7	33.5	35.2	32.4	36.3	32.7
U.S	25.6	36.3	23.9	32.2	23.9	31.3	23.5	29.9
UAE	7.6	2.5	3.2	2.7	3.2	3.0	3.1	2.6
Afghanistan	6.5	0.1	0.2	0.1	0.0	0.0	0.1	0.1
Hong Kong	4.1	5.4	3.9	4.9	4.7	5.7	4.6	5.6
China	2.8	3.6	4.1	3.9	3.7	3.4	3.1	2.9
Saudi Arabia	2.0	1.4	1.3	1.6	1.5	1.8	1.8	2.2
Turkey	1.9	2.4	2.2	2.5	2.1	2.2	2.1	2.2
India	1.8	0.5	0.3	0.4	0.3	0.4	0.3	0.4
Bangladesh	1.6	2.2	2.5	2.1	2.8	2.6	4.3	4.6
Others	20.1	16.5	23.8	16.2	22.5	17.3	20.8	16.8

Source: Federal Bureau of Statistics

countries while remaining 20% is going to the other countries of the globe. Similarly if we observe our total exports in Fy06 we come to know that again almost 80% of our total exports are concentrated in to the top ten countries while only 20% of our total exports are going to the other countries of the world.

Constraints in exports diversification

Shortage of skilled manpower

Skilled man power is instrumental for exports diversification. But in Pakistan due to lack of institutions for technical education we are unable to produce skilled labor force. Our labor force can not utilize the advance technology to produce value added products; as a result we are bound to export only primary products. So in order to diversify our exports we should concentrate on human resource development.

Lack of adequate infrastructure.

We are unable to diversify our exports because of poor economic and physical infrastructure. Our producers are unable to bring the commodities in final markets.

As a result of poor physical infrastructure like roads the farmers especially from the northern areas have to bear high transportation cost. Because of non availability of economic infrastructure our producers are unable to convert low value products into high value added products. Because of poor physical and economic infrastructure there is a lack of coordination between producers and exporters.

Low investment level

Investment is a key for export diversification. But in Pakistan investment level is very low, although there is a rapid increase in investment in the financial and services sector, but there is a lack of investment in the real sector. Due to lack of investment we have less productive capacity which leads to discourage the export diversification.

Under utilization of domestic resources

Although Pakistan has abundant natural resources, but we are not utilizing these resources properly. We have potential in many resources like gems and stones, copper and so on. But due to lack of foreign investment and poor law and order situation we are unable to explore these resources. So for the export diversification it is essential that we should not waste our resources. Proper utilization of resources will lead to enhance the volume of our exports.

High domestic cost of production

High domestic cost of production is another impediment for exports diversification. In order to diversify our exports we need to invest in new machinery and plants as a result our cost of production will be higher at the initial stages. When cost of production is high then we can not compete in the international market.

Lack of marketing

Marketing is another key factor to diversify our exports. But due to lack of proper marketing we are unable to introduce our products in the international market, as a result we are unable explore new directions for our exports.

In effective policy measures

Proper policy formulation is necessary for the exports diversification. But unfortunately in our country whenever we devise a policy our emphasis is to increase the

volume of our exports, but for the export diversification it is necessary that we should formulate a policy
That results into the encouragement of high value added products. in case of Pakistan there are less incentives for the producers to produce high value added products.

CONCLUSION

As the objective of this study was to analyze that whether our exports are diversified or concentrated. Now as a result of above data analysis by using the Lorenz curve and the table. We can conclude that the overall situation is not encouraging. Because our exports are highly concentrated in to the few regions of the world. The most disappointing thing which we can conclude from the above analysis is that no serious effort has been made to diversify our exports since last four to five years we are still stagnant at one point. So it can be concluded that our exports earnings are not stable. Because of high concentration of our exports our export earnings are volatile. In case of any downturn in the political relationships with these few countries will have significant impact on our overall exports. So we should pursue a policy of exports diversification in order to insulate our country from unexpected changes in our terms of trade so that we can stabilize domestic incomes and employment.

Policy recommendations

There are certain policy measures which can be instrumental for the exports diversification. Diversification of exports is helpful for the long term sustainable growth. We need a policy which is fully compatible with challenges of globalization. In this report it is a modest effort to suggest following policy measures.

Education

The structure and quality of vocational and technical education is outdated. And lack effectiveness. Skills aren't developed in proportion to the demand in the job market. So for the exports diversification and to compete in the international market it is necessary that we should educate our labor force. We should develop an education system in which we can ensure the accessibility of the people to technical and vocational education. We should develop institutions which can produce highly skilled labor force.

New products categories

Almost 75% of our exports are concentrated into cotton, rice, leather, and sports. We should direct new products categories like fisheries, poultry, fruit, vegetables, wheat, marble, gems and jewelry, engineering goods, chemicals and general services.

Exploration of new areas

The European Union and North America account for nearly three fifth share in total exports from Pakistan. Our foreign mission and the commerce ministry should play an active role in exploring new markets for Pakistani goods. We have to explore new areas. The EPB has identified Africa, South Africa, Eastern Europe, central Asian republics, and some other areas where Pakistani products can be marketed.

Quality Improvement

For the exports diversification it is necessary that we should improve the quality of our products in order to make them more attractable for the importers. International markets are becoming more and more quality conscious these days. Total quality management (TQM) and continuous improvement (CI) are the order of the day. We have to adhere to these principles if we want to survive in the international market.

Role of the government

The responsibility for the diversification does not rest on the shoulders of exporters only. The government has an important role to play in this regard. The export promotion bureau needs to pursue an aggressive marketing campaign in less explored markets around the globe. Further our trade counselors in such region s should work hard to enhance the country's exports. The government's efforts for free trade agreements with some African countries and srilanka are a step in the right direction.

Bibliography

Massell BF (1964). "Export concentration and fluctuations in export earnings: a cross-section analysis". The American economic review, 54(2: 1): 47-63.
Almas zafar ul Hassan (2003). "Exports diversification and sustainable growth".
Ali R (July 1991). "Is export diversification the best way to achieve export growth and stability?'' 6-7.
Khatoon A (2001) "a strategic approach to exports" 2-3. Daily dawn.
Shehzad I (2003). "Diversifying export markets" 1-2.dawn Karachi. "concentration of exports and imports: an international comparison," Dec. 1958, 68.
Khan AH (2003). "Pakistan's exports: what needs to be done?" 3-4.
Federal bureau of statistics.

Prime Journals of Business Administration and Management (BAM)
ISSN: 2251-1261
Vol. 1(12), pp. 382-391, December 5th, 2011
www.primejournal.org/BAM

to Contents, page 1

Review

The advantages of micro-credit lending programs and the human capabilities approach for women's poverty reduction and increased human rights in Bangladesh

Kazi Abdur Rouf

Social Economic Centre, University of Toronto.

Accepted 3rd December, 2011

Fifty percent of people all over the world are suffering from a lack basic needs. Even poor women are denied their equality of human rights. Equality of rights for these women would mean access to food, clothing, shelter, and credit as well as liberation from exploitative forms of income generation such as domestic work, child labour and trafficking. Women in Bangladesh suffer from inequality of rights on quite an unimaginable level and their socio-economic development has been largely impeded. Thus, these poor women depend on others to survive. Human capability services like education and skills development services are not generated nor tailored to them particularly at the village level which affects the basic human rights of these women. Without educational, health, economic and social services at the village grass-roots level, poor people suffer most in Bangladesh. Although the Constitution of Bangladesh appears to strongly approve gender equality and positive action that guarantees women's full participation in social, economic and political life, it is clear that full support is absent. Ironically, the disadvantaged poor people specifically are struggling to fulfill their basic human needs and are aware of their basic human rights. Although some steps have been made to reduce gender inequality, some laws still lag behind and many discriminatory practices are found in the customary laws, which still remain in force. A solution that has offered to assist poor women in Bangladesh through micro-credit organizations (MFIs) like Grameen Bank credit, that create opportunities for these women to help educate themselves and overcome poverty in Bangladesh. However, human rights education at the grass roots level is very nominal. Hence focus on human rights education extension programs are urgently required to establish basic human rights.

Key words: Grameen Bank, human rights, women human rights, micro-credit, feminization of poverty, participatory development and women empowerment.

INTRODUCTION

The "feminization of poverty" has been a growing concern in low-income countries such as Bangladesh. In this essay, the feminization of poverty refers to the burden placed on poor women in Bangladesh to feed, clothe and nurture their families and themselves. Poor women are denied access to credit. The development of women has been largely impeded, and women in Bangladesh suffer from inequality of rights on quite an unimaginable scale. Equality of rights for these women would mean access to food, clothing, shelter, and credit as well as liberation from exploitative forms of income generation such as domestic work, child labour and trafficking. There is a major lack of resources to support women in their struggle to work, take care of their families and survive. This struggle is often the result of macro- and micro-economic gender divisions, which reflect current political practices and the norms, customs and culture of the patriarchal society. Therefore, the 50% people suffer from a lack of basic needs. Human capability services like education and skills development services are not generated and tailored at the village level which affects basic human rights. Without educational, health, economic and social services at the village grass-roots level, poor people suffer most. Rich people exploit the poor and hinder their human rights and other legal rights. The effects of colonialization and globalization have also contributed to the poor conditions in Bangladesh, first with British rule and then the Pakistan regime between 1947 to 1971.

Although there are several international organizations and programs designed to help combat these injustices, often their techniques are not applied uniformly, and thus civil society has less power to lobby for the necessary changes in the legal system (Quadir, 2003). A promising new strategy, first introduced by Professor Mohamed Yunus, brought to Bangladesh the concept of micro-financing. Micro-credit has proved successful thus far in providing poor women with the ability to generate income while developing human capabilities. The positive effects associated with such activities have helped reduce poverty in some rural areas, as well as increase women's equality and human rights. However, micro-credit and human capabilities development can only thrive in conjunction with successful implementation, government support and services, and international involvement. Thus, this essay will argue that the "feminization of poverty" in Bangladesh can be reduced through the development and successful implementation of micro-credit lending programs that increase human capabilities and provide a means of achieving gender equality and human rights for rural poor women.

Problems

Poor women in Bangladesh do not have basic human rights. Basic human rights in this essay will include food, clothing, shelter, health and education. With no food, clothing and shelter, women are forced to beg, or borrow money. Borrowing money from money lenders leads to great debt, which in turn leads to landlessness and vulnerability in society.

Context

Currently, Bangladesh has 164.4 million people within a 147, 570 square kilometre range. The density of the population is 763 people per square kilometer and the per capita income is $370. In Bangladesh 51% poor people live under the poverty line. This rate is highest in South East Asia. The poverty rate of rural areas in Bangladesh is higher than the national average of 51%. The *Bangladesh Human Development Report 2000* reported that during the period between 1992 and 1996 the rural poverty line declined by about 1% per year and had dropped from 53% to 51%. Yet, a population growth rate of 2.3% per year was not considered in this estimation, thus, despite this decline, rural poverty is still highest in South East Asia.

Because poor Bangladeshi women are not provided state resources, their human rights are seriously hampered, as well as other social civil rights like freedom of choice, freedom of speech, and opportunities in society.

Approximately 86.6% of the population is Muslim. Therefore, Bangladeshi society is a highly patriarchal Muslim dominated society where all household family decisions are made by men. Men control women's labor, women's choice of marriage, access to resources, and legal, social, health, economic and political institutions are mediated by men. Rural women have no autonomy in their life from childhood well into adulthood and old age. Public and private institutions underpin gender subordination and dependence. Women are treated as an unpaid reproductive agent for the family and are involved in unpaid family subsistence agricultural work. Women are discouraged by their male "guardians" not to participate in public or private paid employment outside of the home. Although, public and private sector job opportunities are very limited. Thus, Bangladesh has a high unemployment rate of approximately 52.5%. However, it is important to stress that employment opportunities for women in particular, are required so that unmarried and widowed women can provide for themselves and their families.

Furthermore, there is no value associated with domestic work such as, housework and even social reproduction. Less wage work results in women's low economic status. Poor women do not have access to credit from formal financial services to do business. Some of them become street beggars after losing all that they own through the mortgaging of their assets.

Consequently, these women are left with no rights and no power in the society.

In addition, women who lack education cannot develop the proper skills that will help them find employment. In fact, the adult female literacy rate is 43% (Bangladesh Human Rights Report, 2000). However, the rural female literacy rate is far lower than the national average. If women are unable to find jobs, they cannot provide for themselves and their families. Thus, educated women are more aware of their poor conditions and are better equipped to lobby and fight for their rights.

Only 30% of the total population has access to basic health services, and 76% of all households are deficient in calorie intake (CIDA, 2001). Without access to health care, poor women face malnutrition and death. This makes them very weak, perhaps even too weak to work and provide for themselves and their families. In addition, malnutrition leads to the spread of diseases, which can affect the entire country and not just marginalized sections of the population.

Lack of fundamental resources, affords these women no power and therefore they are unable to exercise their human rights. Kabeer (2003) states that in order to establish gender equality and female empowerment as well as reduce female poverty, Bangladesh needs to close the gender gap in education, increase women's employment and wages, and increase women's participation in parliament.

Causes

Before 1947, Bangladesh was a part of India, and then

to Contents, page 1

between 1947 and 1971 it was under Pakistan. The effects of Pakistan rule in Bangladesh will be discussed later in more detail.

The birth of new social classes (petty bourgeoisie) in Bangladesh before 1947 was the result of the destruction of the old social economy and superimposition of new social policies by the colonial rulers (Alam, 1995). British imperialism created a new class structure, which was absent in India. Although presently it appears that a democratically elected civilian authority governs Bangladesh, it is really governed by the petty bourgeoisies, who are now a ruling class comprised of an emerging middle class, the rich business class, the military and bureaucratic forces.

In addition, there is a history of dependency on foreign capital that exists within the government of colonial powers (Novak, 1993). This government is unable to contribute to a sustainable economic development project for the poor. Therefore, the country is experiencing hegemonic crisis and political instability. Consequently, as Foucault would argue, the notion of power and knowledge related to "participatory development" is inherently anti-nature and anti-women in Bangladesh (Shiva, 1988). As such, although the ruling class in Bangladesh receives foreign aid, the funds are not used for the welfare of poor people. The ruling class would rather use the money for their own interests. This produces misery and inequality in the society (Alam, 1995).

It is clear that British colonialization in Bangladesh has had a negative effect on macro- and micro-economics in terms of gender divisions in labour and within social contexts. These kinds of societal organizations have only supported ideas of subordination and encouraged the birth of many forms of inferiority. Class relations are an important contribution to this discussion because it is often the money lenders that poor women turn to for "support." Poor women in Bangladesh are being exploited by government macro-economic policies and money lenders (*Mohajan*), and therefore are not afforded the basic human rights discussed above. Thus, poor, uneducated women depend on others to survive.

To maintain the family, poor people sell their labour in advance for one, two or even five years. The need to provide food and clothing to children forces poor women to become domestic aids in the homes of rich people. Poor people are forced to move from small rural villages to urban areas for survival. These people live in shanty slums that are in miserably low and unhygienic conditions. A majority of them are suffering from lack of Healthcare. They ultimately end up without food, clothing, shelter, health, education, security, and no choice. They become victims of malnutrition and deadly, crippling diseases. Some brothels take advantage of the troubled times of poor people and begin to traffic and prostitute desperate women and children. The result is an increase of injustice, inequality, and exploitation, which has an even worse impact on their lives.

Another important factor to consider is the maintenance of the male dominated patriarchal society in Bangladesh. Although Bangladesh declared itself a secular state, Muslim religious values dominate the society. These Islamic values are linked with politics in Bangladesh and often, political representatives lobby for their own political interests. Thus, when fundamental Islamic religious leaders become powerful in politics, women's rights are affected. This is because Islam values *Sharia* law which encourages male dominated patriarchy. Parvin Paidar (2002) notes the patriarchal nature of Islam and its subsequent oppression of women in Iran as well. Therefore, in Bangladesh, gender inequality and development discourse needs to acknowledge the contradictory affects of religious customary and legal laws, which are often at the expense of women's human rights.

In addition, as Beneria (2003) notes, with globalization and industrialization, trade became more commercialized, which made profit the motivating factor for work. Market forces in the formal and informal sectors of the economy are also important to the understanding of the feminization of poverty because, as previously discussed; women's work goes largely unnoticed in these sectors as does their unpaid labour and reproductive contributions. Dollar and Gatti (2003) and Seguino (2003) provide some interesting findings about the division of labour throughout their respective studies. These will be discussed later in further detail with a focus on macro- and micro-economic policies and their impact on gender-divisions.

The impact of globalization and commercialism can best be illustrated with a discussion of British Imperialism in Bengal and its post-colonial effects. The British East India Company first introduced British rule in Bengal in 1750. Its purpose was to build a trade market between India and England. Shortly afterward, this company introduced the idea of private ownership similar to the west, and land became an important commodity. A new upper class of "landlords" (Zamindars) emerged who were revenue collectors during the Moghal Emperors. (Karim, 1976). In 1773, the British introduced the Permanent Land Settlement Act. This new system introduced and forced farmers into cash crop production and commercialized agriculture. British metropolises remained linked to India by establishing metropolis satellite structures (Frank, 1970). They diverted agriculture in order to supply raw materials to industries in England, which transformed Bengal into a market for finished products for England. As Boserup's (1970) literature would suggest changes in the production of agriculture effects women and the land, their roles, and the decline of women's equality from pre-colonial to colonial times. Thus, the conflict between gender and class relations becomes more apparent. Although Bangladeshi women contribute to the majority of agricultural production as well as within the informal sectors of the economy, they are still suffering from non-paid and undermined informal household and farming labour.

Another important contribution to the discussion on gender and class relations relates to the British educational policy that created a Hindu educated middle class in Bengal known as “gentlemen” (*bhadralok).* According to Gramsci (1971) this class would fall under the category of “organic intellectuals.” Organic intellectuals are involved in various bureaucratic and professional duties that are necessary to mediate British rule in India. Muslims in India refused to collaborate with the British and rejected learning English as Muslim religious leaders declared the English language the language of “kafirs” (infidels). Therefore, in Bengal, Muslims are behind in education when compared to Hindus. This Hindu-Muslim differentiation later resulted in separate religious sentiments, movements and nationalism to preserve Indian Muslim culture. Thus, in 1947, this division led to the Pakistani acquisition of Bangladesh (East Pakistan).

With new rule, Pakistan introduced development projects which favored West Pakistan’s own interests. This generated class and regional inequalities between West and East Pakistan. Once Bengali petty bourgeois articulated nationalist discourse on the basis of economic and political exploitation, they started to identify political and economic discrimination in the Pakistan public. As a result, in 1971 East Pakistani people revolted, and East Pakistan became Bangladesh.

The development programs in Bangladesh created a parasitic class, who misspent public funds and caused the entire country to become impoverished. Huge foreign aid poured into Bangladesh in the name of public projects, but these funds were mishandled by the corrupt ruling class. This exacerbated the already terrible conditions for poor people.

A third factor contributing to the feminization of poverty today is globalization. Free market capitalism, globalization, democracy and technology are not worthwhile in Bangladesh, and worsen poverty, especially for women. This is because, macro-policies and programs are bureaucratic in nature (top down), and do not consider globalization’s exploitative processes on Third World nations.

The invasion of multinational corporations is similar to the process of colonization. Multinationals come into the country of their choice, and exploit the people, the government and the resources, in the name of business. Like colonization, this changes the role of women in society and affects their human rights. Although the multinational corporations create some jobs in Bangladesh, these corporations care only about profits, which results in uneven economic growth in the country. For instance, in order to increase profits, multinationals find cheap labour and resources. Small businesses fail as they cannot compete with multinational products. In addition, multinational corporations ruin the environment and therefore people have no lands for agriculture, which is needed for survival. As Professor Patricia Stamp urges, we should not pressure nature for human benefit (Stamp, 1989). Human rights are also affected by multinational corporations. People are left to live in poverty, with no options for fighting against multinational corporations. In particular, women have no right to complain about work or complain about poverty and injustice as they have no power. Unfortunately, the Bangladeshi government relies on the money of multinational corporations and so they do not want to be rid of them.

Globalization literature (also referred to as “modernization literature”) ignores traditional values and Third World countries are often depicted as falling behind. Professor Patricia Stamp disagrees with such literature and argues that it is not a question of whether such Third World societies are “behind,” but rather, modernization imposes Westernized views that are not tailored to consider the economic, political and social practices of traditional Third World societies. Edward W. Said (1978) has noted that “orientalism” (the erroneous Western tendency to view all Middle Eastern and Asian cultures as a kind of homogenous whole and to place them within a false framework that is opposed to our own) is man-made for the West. The relationship between the Occident (Britain and the U.S.) and the Orient (Middle East and East Asia) is a relationship of power, a domination of hegemony. The interaction between the ruling classes and the state is the end result of a historical process to keep the country undeveloped (Said, 1978). This is very true as we note the effects of western capitalist goals with respect to modernization.

Diane Elson (2003) provides some insight into becoming gender-aware in macro-economic policies and government budgets. She notes the importance of unpaid reproductive labour and the dynamic between gender inequality and capitalist markets. Macro-economic policy developments need to be evaluated and adjusted to reorganize gender differences with respect to economic growth.

Attempted solutions

The Bangladeshi Constitution calls for equality of all citizens before the law and no discrimination against any citizens on grounds of religion, race, sex, or place of birth; right to protection of the law; protection of right to life and personal liberty; freedom of speech, profession or occupation; rights to property; and enforcement of fundamental rights through courts of law. Any laws and enactments inconsistent with fundamental rights are void. The constitution of Bangladesh guarantees equal rights for women and men. However, there is still a large gap between the law and its actual implementation.

Article 25 of the Bangladeshi Constitution is vital to this discussion because it addresses the fulfillment of basic

to Contents, page 1

needs. Everyone has the right to a standard of living, adequate for the health and well-being of himself and of his family, including food, clothing, housing, medical care and necessary social services. However, in the constitution a provision for the same is made in Article 15 that the fulfillment of these basic needs is not a matter of rights, but rather, a state responsibility, which was unfortunately not included in the fundamental rights chapter of the Constitution. The above points illustrate that the constitution provides for the fundamental rights of its citizens, but fundamental needs are not fully recognized as rights of citizens. In addition, Article 10 of the Constitution notes, "Steps shall be taken to ensure participation of women in all spheres of national life." (BMSP, 1997). Ironically, the bottom 50% of disadvantaged poor people specifically is struggling to fulfill their basic human needs.

Although the Constitution of Bangladesh appears to strongly approve gender equality and positive action that guarantees women's full participation in the social, economic and political life, it is clear that full support is absent. There is a separate ministry for women's affairs and although some steps have been made to reduce gender inequality, some laws still lag behind and many discriminatory practices are found in the customary laws, which still remain in force.

A solution that was offered to assist poor women in Bangladesh began with Professor M. Yunus. Professor Yunus is the founder of the Grameen Bank. This organization argues that welfare or handouts do not help poor people (Yunus, 2002). Instead, the poor remain unskilled and continue to live in poverty. Grameen Bank credit creates opportunities for poor women in Bangladesh to help poor women educate themselves and overcome poverty.

Grameen Bank, a locally initiated model, provides credit to rural poor women without collateral. They serve 4 million families and provided 4.5 billion US dollars to its borrowers across Bangladesh. The credit recovery rate is 99% (Grameen Bank, 2005). It has not only had tremendous success in generating income to the bottom 50% disadvantaged women, but it has also empowered them to make choices, have a voice, and gain opportunities and bargaining power. It views credit for self-employment as a fundamental human right, which is a powerful weapon that grants access to other resources. Grameen Bank has helped poor women break out of the cycle of poverty by increasing the income of its borrowers. Grameen clients are able to overcome the deprivation of basic human needs and fundamental human rights.

Following the Gramen Bank model, other micro-credit organizations (Bangladesh Rural Advancement Committee, Association for Social Advancement, PRISHIKA, PKSF) in Bangladesh are also now providing credit to poor village women for income generating activities. Various studies in Bangladesh show that credit programs have a positive impact on the reduction of poverty. Kofi Annan (2005), the former Secretary General of the United Nations, said on January 15, 2005 that micro finance has proved its value in many countries as a weapon against poverty and hunger. It really improves peoples' lives, especially the lives of those who need it most (Annan, 2005). Therefore, it is very important for poor women to have an economic income base that can open the door to other social and political rights for women, equality, freedoms, as well as struggle against the violations of their human rights and social injustice.

International organizations also play a vital role in helping to reduce the feminization of poverty. One such organization is the United Nations (UN). The United Nations has helped the world become more aware of this problem. A new strategy has been proposed by the UN that will include trying to implement development programs through the government. For instance, the Millennium Development Goals (MDGs) were agreed upon at the United Nations Millennium Summit in 2000 and include: (1) to half world poverty by 2015; (2) to achieve universal primary education and (3) to promote gender equality in order to help empower women. However, as previously discussed, some corrupt government bodies divert funds for their own preferences. Furthermore, rural poor people are not considered a priority in the government fiscal budget agenda. Poor people are suffering from the disadvantages of national budget resources. However, Women's equal access to financial resources is a human rights issue.

Other international organizations include, The World Bank, International Monetary Fund (IMF) and the World Trade Organization (WTO). These organizations took Third World countries and tried to push structural adjustment policies (SAPs) without tailoring the SAPs to each country. SAPs were intended to help the economic disturbances in Third World countries like Bangladesh. However, foreign debt adjustment did not consider gendered responsibilities in and outside the home. The World Bank promotes economic progress in developing countries. They try to raise productivity to help people live better lives. Their priorities are at the macro-level, but have less impact at the micro-level. IMF monitors international finance, encourages financial cooperation between countries, lobbies for state exchange rates and assists governments with debt. However, IMF does not specifically address women's issues. It does not focus on poverty reduction and its main concern is economic growth. Equal economic growth is also not addressed. IMF does not provide hands–on assistance. Their SAPs are designed only to correct maladjustments in their balance of payments. The WTO was formed in 1995 and its basic principle is that all 125 member countries must abide by WTO rules. Rules include barriers on tariffs, intellectual property rights, and investment and trade relations. However, the WTO does not address worker's rights and poverty.

A proposal offered by the World Bank and the IMF was the greater funding and implementation of non-governmental organizations (NGOs) through apex funding. The involvement of apex funding was required in order to help monitor the activities of NGO's. Otherwise, there would be an overlap of resources in some areas, leaving other areas with nothing. The autonomous apex funding body in Bangladesh is called the Pally-Karmar-Foundation (PKSF). They take funding and distribute it to NGO's and coordinate and monitor NGO activities. The advantages of this model include, less government control, and funds go directly to NGOs. NGOs provide services directly to the people. Some NGOs are working for mass education, some are involved in primary health care services, some provide agricultural support services and some promote women's human rights. The majority of NGOs are now involved in micro-credit programs, which target poor rural women and promote income-generating activities. Through diversified NGO activities, village people are connected to various development programs. The Ain-O-Salish Kendra undertakes a variety of programs to improve people's awareness of legal and human rights. Mohila Samity (Women Association) runs various programs to help women achieve self-reliance. In particular, in Bangladesh, many NGO's work with women; this is a massive thrust toward women's equality and progress.

Two international organizations born in the west include the Canadian International Development Agency (CIDA) and the US Agency for International Development (USAID). CIDA supports sustainable development in developing countries. They hope to help reduce poverty and make the lives of poor people more secure and equitable. CIDA also supports democratic development. Its program is based on the Millennium Development Goals (MDGs) developed by the United Nations. CIDA has 4 main priorities: 1. social development; 2. economic well-being; 3. environmental sustainability; and 4. governance. There is a greater emphasis on human rights issues, democracy and good governance to help reduce corruption by governments. Also, equitability between men and women is promoted and supported by CIDA. The CIDA approach is more appropriate for Bangladesh in order to increase gender equality, human rights and human capital development of poor women.

On the other hand, US foreign assistance has always had the twofold purpose of furthering America's foreign policy interests in expanding democracy and free markets while improving the lives of the citizens of the developing world. USAID appears to be very general and vague about their policies and work in developing countries. Their policy objectives focus on economic growth, trade and democracy, and conflict prevention. USAID is unclear about how it helps economic growth and eradicates poverty. They do not focus on women's human rights and equity. They focus on terrorism and weapons of mass destruction. Oddly, education is not a priority even though education is a basic requirement for human capital and for human rights. Thus, it is clear that CIDA support is more oriented to promoting human rights, reducing poverty, and macro- and micro-level social and economic development in developing countries. Conversely, USAID is more concerned with US interests like foreign policy, commercialism, trade, globalization and an emphasis on capitalist democracy within a US context.

The purpose of these organizations is to stabilize currency and ensure free trade. Yet, none of these organizations have yet emerged as key contributors in Bangladesh civil society (Quadir, 2003). The big three (World Bank, IMF and WTO) reflect the interests of the capitalist powers. In fact, there is public opposition against some World Bank policies. Not only do people feel that these institutions develop devastating SAPs which negatively affect the poor, but feminist scholars also believe that the World Bank ignores women's issues and maintains gender inequality. Currently, a piecemeal system exists between NGOs and civil society. They need more coordination so that a complete package is implemented at the grass-roots level. Therefore, Bangladesh needs a national NGO who solely deals with the village people to counsel, advocate and lobby for women's human rights, stand against dehumanization, and work for victims assisting them in legal issues and prosecution.

Directions for the future

As public and private sector formal jobs are limited, government resources and support for the villages is bleak. Rural poverty is increasing at an alarming rate in comparison to its population growth. Therefore an independent income program, a universal education program and strong law enforcement can create the potential to enhance women's income, awareness, capabilities and status at home and in the community. Income and gender specific programs can benefit women because "micro-finance is a vital means for income generation, social inclusion, and empowerment" (Chowdhuray, 2005). The Grameen Bank program has proved the success of micro-credit. Distressed women are provided with education, skills, training, credit, and other support services for income generating purposes. This leads to economic progress, a boost in family and social status, independent decision-making, and the development of self-confidence that empowers. Women's empowerment may lead to government participation at the local level. This participation is important for achieving the goals associated with women's human rights.

Increased human capabilities have a positive impact on both gender equality and economic growth (Nassubum, 1988). Studies conducted by Dollar and Gatti (2003) and Seguino (2003) illustrate a positive relationship between

to Contents, page 1

gender equality, efficiency and economic growth. In addition, Seguino (2003) suggests that there is a trade-off between gender equality and economic growth. Therefore, investment in human capital will improve efficiency and will inevitably have a positive impact on economic growth.

Micro-level programs, such as micro-credit, will help eradicate social discrepancy, gender inequalities, and other social catastrophic activities, which affect the poor women in rural Bangladesh. However, programs such as Grameen Bank (GB), along with other micro-finance institutions, only serve 14 million rural people composed of GB 4.00 million, PKSF 5.2 million, Bangladesh Rural Advancement Committee (BRAC) 3.9 million, and Others 1.00 million and therefore, a large population still has no access to financial services. Therefore commercial banks and other financial institutions need to be government regulated to simplify collateral conditions and reduce banking bureaucracy.

Raising awareness through legal education and reform can challenge existing public policies and laws to empower victimized women. As Paulo Freire (1981) would suggest, ignorance leads to unequal power and exploitation. Thus, the government should promote and execute universal compulsory primary education with a special emphasis on female education. This will help poor women develop human capabilities and human dignity.

Besides education, the government needs to initiate greater support at the social, economic, and political micro-level. For instance, rural poor women are involved in informal agricultural and non-agricultural work. Their work is not included in the country's economic production and social reproduction goes unnoticed. Because of the domestic (household) nature of this work, it is stereotyped as "women's work" which leaves poor women out of the fiscal budget. Therefore, the government should develop labour intensive special income generating projects in the rural areas for these poor women so that their activities will be included within the development programs.

In addition, deficient macro-policies can easily limit the development of poor women. A micro-economic policy framework for maintaining poor women's development may be more conducive to sustainable development and poverty reduction. Social mobilization is a pre-condition for improving the access, equity and quality of public services available at the local level. The government of Bangladesh can implement a massive social mobilization program to encourage rural communities to solve their own common concerns and share any available local resources among themselves.

A special judiciary bench may be established at the district level to deal with violence against rural poor women and to enforce laws regarding trafficking, dowry, rape, forced prostitution, and acid throwing. Courts at the local level need to criminalize and penalize the traffickers, syndicates and operators. At the same time, the victims of these crimes need to be empowered in order to help them rebuild their lives. This can be accomplished through rehabilitation in health and the provision of other social services. Men should be made to take responsibility for their sexual behavior and so the courts need to strictly enforce the existing laws with the help of honest law-enforcing personnel.

Margaret Shuler's "Strategy Matrix" provides techniques to work with the structural components of the law, so that we may find a strategy to assist women in Bangladesh who face violations of human rights. Strategies need to improve women's access to the legal system and services. The gender network committee at the upzilla (sub-district level) can be set up to monitor legal actions and enforcement of legal laws for establishing women's human rights.

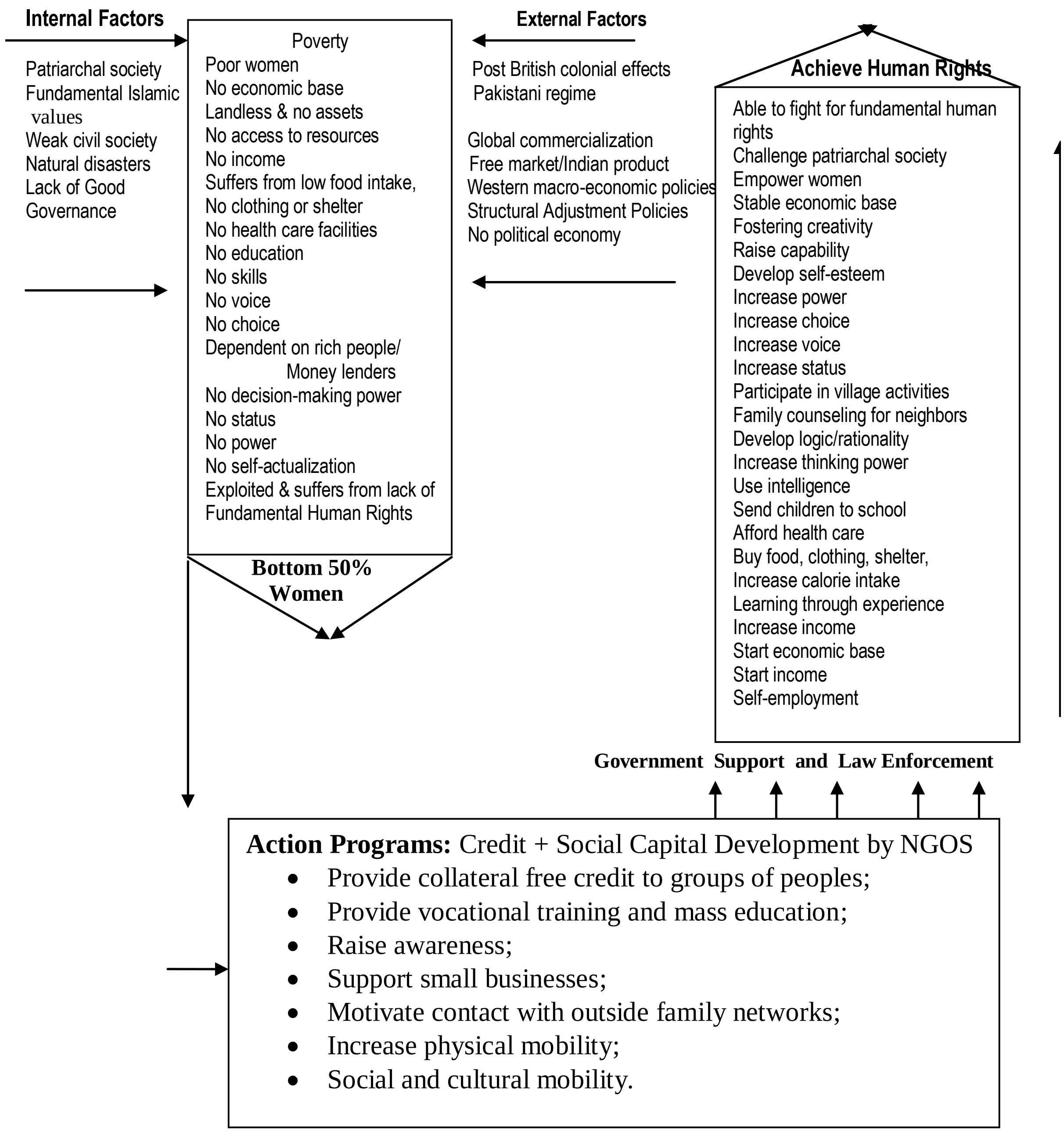
At the same time the international community should also come forward and create pressure for governments to take serious action and to eliminate poor women's human rights violations. The Millennium Development Goals are an example of such action. However, without the development of women, Bangladesh will not be able to successfully accomplish their MDGs. Therefore, if a country wants to address the feminization of poverty, it also needs a clear definition of who constitutes poor women (bottom 50%) so that any projects or programs designed to assist poor women can be exclusively created for the bottom 50% of women. Figure 1 summaries a human right framework for the empowered of marginalized women in Bangladesh.

Although this essay covers all three MDGs as well as their targets and indicators, I want to stress the urgent need to include another target under the first MDG. That is to establish poor women's human rights by promoting income generating programs and self-employment through the provision of collateral free micro-credit. The urgency of this proposition is important as poor women often go unnoticed in the development agenda, as they are invisible in society (Hick, 2004). For instance, the MDG that promotes universal primary education is too general, and does not specifically target poor women. Without being clearly targeted, women will be left out of education because of extreme poverty. Although the Bangladesh government has a good policy with respect to increasing literacy rates, this policy will only reach the poor with good governance and strong support for such national strategies. Good governance and government support is essential for ensuring that poor women are enrolling in vocational training and technology programs and continuing universal education.

CONCLUSION

The researcher believes that an independent income program can create the potential to enhance women's status at home and in the community. The Grameen Bank program has proved this through its micro-lending program. As Kabeer (2003) notes, by increasing women's

Figure 1: Human Rights Diagram for Bottom 50% Poor Women in Bangladesh.



income, women can have an increased number of choices, networking, greater household decision- making power, greater social status, and greater sense of confidence and independence. Women's empowerment can enable women to oppose authoritative patriarchal power structures through collective action.

Empowering women through education is a big key to change. Otherwise, their human capabilities will not develop, and rich people will eat the fruits of the economy while exploiting and abandoning the poor people. Therefore, there must be opportunities and choice provided for them. This will help to develop their human capabilities and help reduce poverty among women.

Other social, economic and political support at the micro level can assist poor women in the development of their human capabilities. For instance, the rural population needs to operate a variety of self-employment activities. Access to capital via diverse micro-credit and micro-enterprise targeted loan schemes helps to remove the credit constraint and hence accelerates the exits out of poverty. Grameen Bank is proofs that micro-credit can create economic and social capital among the rural poor

to Contents, page 1

women in Bangladesh (Dowla, 2001). These programs can help to eradicate social discrepancy, gender inequalities, and other social catastrophic activities that affect the bottom 50% of women in rural Bangladesh. Figure 1.

Women's human rights development discourse differs between developed and developing countries. There are various approaches discussed by different authors within the literature. The literature helps us to understand problems and the position of women in Third World countries. However, I am inclined to agree with Patricia Stamp (1989) and Chandra Mohanty (2003) when they stress the importance of analyzing the use of the literature in development discourse. More specifically, the literature should be used to promote poor women's human rights in Bangladesh, but in such a way that accepts that not all the traditional values of Bangladesh are a constraint to development discourse and practice, and not all Western ideologies can be applied to Bangladesh without considering local contexts. As Beneria (1970) would suggest, development relates to Third World countries, gender and globalization, economic adjustment and feminist economics. These themes should be examined as interrelated dimensions that are relevant to external factors affecting the Bangladesh poverty process and inequality. Freire (1981) notes that through constant dialogue and praxis we can free each other and free ourselves from oppressors. Therefore, people must work together to a find permanent freedom.

Margaret Schuler (1986) offers us a method for understanding and challenging the current legal system. Her concept of the structural component of society (that is, courts, administrations and law enforcement) should be challenged through advocacy. Building networks and organizing public and private protests and rallies can help to fight against women's exploitation. In addition, activities such as seminars, conferences, workshops, community education, mass media campaigns, and publication of scholarly works, dissemination of information through popular literature, comic books, posters, dance, brochures, theatre, and poetry can initiate programs for women designed to help them claim their inherent rights.

Furthermore, the laws should be reformed in order to allow women freedom to claim their property rights, and freedom of choice. At the same time, the law needs to raise awareness through legal education to reform unequal power for social justice in the society. Greater government intervention and less government corruption is also needed to assist these women.

All over the world, poor women are living such inhuman lives and are exploited by government and the powerful elites. Therefore, it is urgent that we work to establish basic human rights for the poor women. It becomes a question of morality and thus, our duty to help these disadvantaged people survive in society. Empowerment and autonomy of women and improvement of women's social, economic and political status is essential for the achievement of both a transparent and accountable government, as well as sustainable women human rights in Bangladesh.

ACRONYMS

Bangladesh Rural Advancement Committee (BRAC), Canadian International Development Agency (CIDA), Grameen Bank (GB), micro-finance institutions (MFIs) Miullinium Development Goals, (MDGs), non-governmental organizations (NGOs), Palli Karma Shayak Foundation (PKSF), Structural Adjustment Policy (SAP), United Nations (UN), World Trade Organization (WTO) and World Health Organization (WHO).

ABOUT THE AUTHOR

Kazi Rouf, PhD in Adult Education and Community Development at the Ontario Institute for Studies in Education (OISE). He did his Masters in Environmental Studies, Business and Environment (MES) from York University in 2008 and Bachelor of Arts (BA) in Women Studies from York University in 2006. Rouf has worked for Grameen Bank Bangladesh and in other micro-finance institutions in Bangladesh, India, Afghanistan, and Pakistan. Moreover, he worked in Namibia, Lesotho, India, Pakistan, Afghanistan, Philippines, United States and Canada for promoting small and medium enterprises (SME) and self-employment among low income people. Kazi Rouf writes many articles after working for Grameen Bank and other micro-finance institutions around the world for long time. To contact Kazi and for details regarding the references contained in the article please contact him at: rouf56@hotmail.com.

REFERENCES

Alam SMS (1995). *The state, class formation, and development in Bangladesh*. New York: University Press of America, Inc.

Annan K (2005). *Grameen dialogue #60, January 2005*. Dhaka, Bangladesh: Grameen Trust.

Beneria L (2003). *Gender, development and globalization: Economics as if people mattered.* New York: Routledge.

BIDS (2005). Fighting human poverty: Bangladesh human development report 2000. An overview of the study prepared by the *Bangladesh Institute of Development Studies (BIDS)*. Retrieved June 17, 2005, from http://www.sdnbd.org/sdi/issues/sustainable_development/bd

BMSP (1997). *State of human rights Bangladesh.* Dhaka, Bangladesh: Manobadhikar Samonoy Parishad (BMSP).

Boserup E (1970). *Women's role in economic development*. London: Earthscan Publications.

Bangladesh Rural Advancement Committee (BRAC). (2005). *BRAC information*. Retrieved July 06, 2005, from http://www.brac.net/

Canadian International Development Agency (CIDA). (n.d.). *Development goals.* Ottawa: International Development Research Center. Retrieved June 5, 2005 from http://www.acdi-cida.gc.ca/whatwedo.htm.

CIDA (2001).*Gender profile: Bangladesh*. Retrieved June 5, 2005, from http://www.acdi-cida.gc.ca/cida_ind.nsf/0/60/3ea7adbb8d592e85256b1b007403ef?OpenDocument#1

Chouwdhury A (2005). *Grameen Connections*. Newsletter, *7(4).* Washington, USA. The Grameen Foundation

Dollar D, Gatti R (2003). In N. Kabeer, *Gender mainstreaming in poverty eradication and the millennium development goals* (pp. 39-44). Ottawa: International Development Research Center.

Dowla A (2001). *Grameen Dialogue #60, January 2005*. Dhaka: Grameen Trust.

Elson D (1998). Talking to the boys: Gender and economic growth models. In C. Jackson R. Pearson (Eds.), *Feminist vision of development: Gender analysis and policy*, (pp. 155-170). New York: Routledge.

Elson, D. (2003). Gender justice human rights, and neo-liberal economics politics, In Molyneux, M. & Razavi, S. (Eds.), *Gender justice development, and rights* (pp 78-114). New York: Oxford University Press.

Frank AG (1970). Dependence, imperialism and underdevelopment. In J. Cockcroft, A. Gunder Frank & D. Johnson (Eds.), *Dependence and Underdevelopment: Latin America's Political Economy*, (pp. 3-17). New York: Doubleday & Company, Inc.

Freire P (1981). *Pedagogy of the oppressed.* New York: Continuum.

Grameen Ban. (2005). *Grameen Dialogue #60, January 2005*. Dhaka, Bangladesh: Grameen Trust.

Gramsci A (1971). The Intellectuals. In Q. Hoare & G. N. Smith (Eds.), *Selections from the prison notebook*, (pp. 4-23). New York: International Publishers.

Hick S (2004). *Social welfare in Canada-understanding income security*. Toronto: Thomson Educational Publishing, Inc. International Monetary Fund (IMF). (n.d.). Retrieved June 17, 2005, from http://www.imf.org/external/pubs/ft/aa/aao1.htm

Kabeer N (2003) *Gender mainstreaming in poverty eradication and the millennium development goals.* Ottawa: International Development Research Center.

Karim A, Nazmul K (1976). *Changing society of India, Pakistan and Bangladesh.* Dhaka: Nawroze Ketabistan.

Mohanty C (2003). *Families without borders: Decolonizing theory, practicing solidarity*. Durham: Duke University Press

Nussabaum M (1988) Women's capabilities and social justice. In Molyneux, M. & Razavi S (Eds.), *Gender, justice, development and rights* (pp 45-74). New York: Oxford University Press.

Natsios AS (2005, April 22). Democratic opportunity and the Islamic World. Speech presented by Andrew S. Natsios on behalf of the *Centre for the Study of Islam and Democracy*. Retrieved June 2, 2005, from http://www.usaid.gov./press/speches/2005/sp050422.html

Novak JJ (1993). *Bangladesh reflection on the water*. Bloormington: Indiana University Press.

Paider P (2002) Encounters between feminism, democracy and reformism in contemporary Iran. In Molyneux, M. & Razavi, S. (Eds.), *Gender Justice Development, and Rights,* (pp 78-114). New York: Oxford University Press.

Pally Karma- Sahayak Foundation (PKSF) (2005). *PKSF profile*. Retrieved July 07, 2005 from http://www.pksf-bd.org/aprm_summit_new_list_workshops_panelists.htm

Quadir F (2003). How "civil" is civil society? Authoritarian state partisan civil society and the struggle for democratic # development in Bangladesh. *Canadian Journal of Development Studies*, *24(3),* 425-438.

Said EW (1978). *Orientalism.* New York: Penguin Books.

Schuler M (Ed.). (1986). *Empowerment, and the law: strategies of Third World women*. Washington, D.C.: OEF International.

Seguino S (2003). In N. Kabeer, *Gender mainstreaming in poverty eradication and the millennium development goals* (pp. 39-44). Ottawa: International Development Research Center.

Shiva V (1988). *Staying alive: Women, ecology, development*. London: Zed Press.

South Asia Foundation. (2001). *Bangladesh profile*. Retrieved on June 08, 2005, from http://www.southasiafoundation.org/saf/bngladesh/banglastats.asp

Stamp P (1989). Technology, gender, power in Africa. Ottawa: International Development Research Centre.

United Nations. Department of Public Information. (2000). Gender Equality, Development and Peace for the Twenty First Century. *Women 2000: Beijing '00.*

Proceedings of a conference on Women's rights. New York. 5-9 June 2000. Retrieved June 17, 2005 from, http://www.un.org/womenwatch/daw/followup/session/presskit/fs1.htm

US Agency for International Development (USAID). (n.d.). Retrieved June 17, 2005 from http:///www.usaid.govt/about_usaid/. World Bank. (n.d.). *Data and statistics for Bangladesh.* Retrieved June 17, 2005, from http://www.worldbank.org/

World Trade Organization (WTO). (n.d.). The future of the WTO. Retrieved June 17, 2005, from http://www.wto.org.

Yunus, M. (2002). *Grameen Dialogue #52, October 2002*. Dhaka, Bangladesh: Grameen Trust.

Prime Journals of Business Administration and Management (BAM)
ISSN: 2251-1261
Vol. 1(12), pp. 369-375, December 5th, 2011
www.primejournal.org/BAM

to Contents, page 1

Review

Risk management and corporate governance in Nigerian banks: A review

[1]SO Uremadu* and [2]AM Azeez

[1]Associate Professor, Department Of Banking and Finance, CABFM, MOUAU, Abia State
[2]Department of Accounting and Economics, Faculty of Business Studies, University of Oxford, UK Great Britain

Accepted 13th November, 2011

In this paper the concept of risk management and its applicability to business operations in Nigeria with a view to raising the banking firm's performance was brought into sharp focus. We employed three research methodology approaches: descriptive, industry and economic analysis models in the study to investigate the applicability or otherwise of risk management techniques and corporate governance principles to raise banking sector's performance in Nigeria. Findings showed that (1) Nigerian banks indulged in excessive risk taking and never cared to apply risk management techniques over the years thereby leading to total collapse of majority of them, (2) Causes of bank failures arising from poor risk management are giving loans without proper appraisal and documentations, giving loans to public (political) office holders with their positions (status) as collaterals, poor internal control system, insider abuses, poor corporate governance practice, financing illicit contracts, etc and (3) Banks carry various kinds of risks in their normal business of banking such as lending and borrowing and other financial services activities. We therefore recommend that if full application of risk management techniques and concepts embedded in the basic principles of corporate governance and revival of culture of trust in our society and among bankers and their customers are brought to bear (applied) in the business of banking in Nigeria, the banking sector's performance will greatly improve thereby raising the economic growth of Nigeria in the years ahead. If our recommendations in the paper are applied to the letter they will reduce, to the barest minimum, corruption and lack of application of risk management techniques and corporate governance in the Nigerian business environment, so to speak.

Key words: Risk management, corporate governance principles, business environment, Nigeria.

INTRODUCTION

Profit maximization or rather maximization of shareholder wealth has been regarded as the main driving force of any business organization (Philippatos, 1974). A firm with good management will certainly make continuous efforts to ensure that it chooses the right combination of investments, financing as well as dividend policy decisions that will maximize its value to its shareholders (Azeez, 2010). This drive affects its expected risk-return characteristic. Given the fact that the return to stakeholders is not known with certainty, risk is always imminent.

Van Horne (2001) views risk as the possibility that actual return will deviate from that which was expected. Expectations are continually revised on the basis of new information about investment, financing and dividend decisions of the firm. In other words, on the basis of information about these three decisions, investors formulate expectations as to the return and risk involved in holding a particular stock.

Rational investors have the habit of seeking to know what the compensation for an investment is likely to be before they put their capital in that investment. The expected return is termed as return. Risk arises from the uncertainty of the actual return which may not be the same as the expected return.

A lot of factors contribute to this deviation from expected returns which include change in government or fiscal policies that can result to interest rate going high or low, increase in taxes, advances in technology, threats of competitors, terrorism threats and so forth. As a result of these challenges, business firms face one form of risk or the other in their operations (Azeez, 2010).

to Contents, page 1

However, where these risks are not properly managed, there is every possibility of the investors' income or capital being lost or even in some cases both are completely erroded. To cover or tackle this uncertainable there is need for investors to devise risk management techniques in order to effectively manage the risks that they face in the day to day operations of their businesses. As once said by Carter (1995) in Azeez (2010) "we bankers must develop cultures that foster and reward the management of risk. We must continually update our risk management policies to ensure that they reflect changing industry dynamics".

According to Dickson (1991), the concept of risk management requires the identification, evaluation and economic control of those risks, which threaten the assets or earnings capabilities of a firm. The adoption of risk management techniques (or models) makes it possible for the exposure of the business to different risks to be managed properly. This means that given certain promised return, a less risky investment will be favoured over the more risky investment. In this dynamic and unstable global business environment of today, investors may not be advised to put risk management techniques in place (more especially in most developing economics of the world like Nigeria where not many companies have risk management techniques in place) to ensure the survival of their investment and returns (see Azeez).

Statement of the problem and the research questions

All kinds of businesses whether banking finance, retail, construction, oil and gas, name it, are surrounded by one form or risk on the other. In other words, business firms cannot exist without some form of risk surrounding them. However, not all businesses put risk management techniques in place despite all the risks they are surrounded with more especially in the Nigerian business environment necessitating this study. The investors or business managers are not proactive enough in other words, they do not plan against these risks, they just wait for the eventualities to happen then they start contemplating on how to overcome the eventualities. This is fire-fighting kind of management style as put forward in (Hicks and Gullett, 1981).

Therefore, the hypothesis proposed for this study is: Risk management is significantly relevant to business operations in Nigeria like the banking firms. The reason is that not all business firms see risk management as one of the most important functions or objectives to be pursued in the course of piloting the affairs of their companies even the big players in the economy like banks and telecommunication companies are in this category, because of this the Central Bank has put practice of risk management and corporate governance to be a must but not all have been abiding by it. Yet they have not been complying with the regulation in some like the credit risk issue in which the government had to bail out eight commercial banks such as Oceanic Bank of Nigeria Plc, Intercontinental Bank Plc, and others in 2009 during the global financial crisis and just recently, three of the former eight troubled banks again had to be rescued from total failure through an abridged bank model arranged by the Central Bank of Nigeria (CBN), to redeem them, to be specific.

Three questions will therefore need to be answered by this study:

1. Has risk management and corporate governance any relevance in managing business operations (including the banking firms) in Nigeria?
2. Why do business firms need to devise risk management techniques usually encapsuled in the principles of corporate governance in managing Nigerian firms (including the banks)?
3. Can risk management practices be incorporated into corporate governance principles in the process of managing Nigerian firms (including the banks) and how?

By the time these questions raised in this study are critically analyzed and tackled, answers to them would have led to discover as to how to tackle the issue of relevance of risk management in the Nigerian business operations in general, and in the banking sector, in particular.

Objectives of the study

The major objectives articulated for this study aim to achieve are:

1. To identify the relevance of risk management in business firms by identifying the various classes of risks likely to be faced by Nigerian businesses, their consequences on income and capital if not managed and then make recommendations on the application of risk management concept or model if found useful. This will also be based on enshrining principles of corporate governance in all our business operations practice with the banking firm inclusive.

METHODOLOGY

In this paper, we adopted three approaches employed by Uremadu (2005)'s work: the descriptive industry and economic analysis models to discuss meaning, nature, scope and features of risk management (embedded in corporate governance principles) as applicable to the business of banking in Nigeria, in particular and globally, in general. Study also utilized Azeez (2010)'s data to pursue a comprehensive analysis of the work in hand especially, from the explorative angle of the discourse. Specifically, paper gave a basic definition and meaning of risk management and as it applies to the practice of banking in the financial services sector of the Nigerian economy, reviewed relevant body of literature on the concept of risk management in the banking industry as well as different methods of risk management in commercial banks to obtain an effective and efficient operations that will lead to an increased banking performance depicted in high expected returns and or higher profitability. It also centered discussions around core risks surrounding the financial institutions, established findings from the study and proffered solutions that will lead to an improved banking system in the Nigerian economy via proper application of risk management principles to the increased growth of the domestic economy via better performance of Nigerian banks in the 21st century world of our time.

LITERATURE REVIEW

Here we exploratively and analytically review a body of literature relevant to the study.

General descriptive, economic and industry Analysis of risk management in commercial banks: A review

In all human endeavours not only business, one thing that is constant is risk for there is no certainty of success in that endeavour. Different scholars and writers have given their views/definitions of what risk is; in other words, we have no one single definition for risk. According to the Concise Oxford Dictionary in Azeez (2010), risk is "the chance or possibility of damage, loss, injury or other adverse consequences". According to Sanusi (2007), risk can be defined as the probability that outcomes vary from our expectations. He further argued that it is the threat or possibility that an action or event will adversely or beneficially affect firm's ability to achieve its objectives. Risk is exposure to uncertainty of outcome (Cade 1999).

In general, risk entails two essential components, which are exposure and uncertainty. Risk then is exposure to a proposition of which one is uncertain. In the case of a business firm, the presence of risk brings about the uncertainty of profit and or success in every investment. On the other hand, uncertainty as argued by Cade (1999) can also be reflected in the volatility of potential outcomes plotted on a probability distribution curve, for which the normal measure of dispersion could be either the variance or the standard deviation. In this case, the wider the standard deviation, the greater the volatility; and thus, in theory, the greater the uncertainty or risk, so to say.

Uncertainty based on a common usage is a state of not knowing whether a proposition is true or false. In this case, probability is often used as a metric of uncertainty. Still, according to Cade (1999), outcome is the consequence of a particular course of action.

For a business to succeed or survive products must be sold and debts must be collected natural hazards such as floods, earthquakes, etc. are there to be confronted, so also corporate governance issues, corruption, terrorism acts, systems failures, etc are to be considered too. On television and radios, one frequently hears these kinds of events happening almost on a daily basis more especially in places such as Nigeria where there are weak laws and regulations. It is also common to see headlines carrying about news of robberies, oil spillage, kidnapping of foreign workers in the Niger-Delta region of Nigeria, and so many other similar incidents of like fashion. Therefore, it has become imminent for organizations more especially the banking firm or industry that by nature is very debt massive in its capital structure (Uremadu, 2009) to prepare for such events which if (they) happen could hinder the normal operations of the business. This can only be achieved by putting in place contingency plans such as devising risk management techniques in every operation of the banking business or firm. As has already been mentioned earlier, risk is not always bad. It was Suzanne Labarge in Azeez (2010) who argues that risk is mismanaged, misunderstood, mispriced or unintended. Whereas, Cade (2007) emphasized that risk is to be respected not shunned. He went further to argue that we enterprise can achieve anything without engaging in risk and the business of banking is characterized by the way in which it underpins the financial risks of the community too often it must be admitted, at an adequate rate of return. The globalization and for the fact that the banking firm's capital structure is debt massive and so forth (Uremadu, 2009 and Azeez, 2010). These kinds of risks include, credit risk, liquidity risk, political risk, market risk, operational risk, interest rate risk, amongst others. These components of risk variables will be considered, in detail, on a future study that will involve determining factors influencing risk management among the Nigerian banks. Writers generally identify between three and ten basic risk categories, depending on which they consider being primary and which they consider to be secondary which they are similar but the emphases often differ (Azeez, 2010).

According to Angelopoulos and Mourdoukoutas (2001) in Azeez (2011), the risks that surround the financial institutions can be categorized into two main categories, which are; traditional and non-traditional categories.

(i) Traditional financial risks are risks arising from the basic function of banks and their intermediaries, that is, as borrowers, lenders and inventors funds. They went further to argue that traditional banking risks can be classified into four categories, namely: liquidity risk, credit risk, political and legal risk and operational risk.

(ii) Non-traditional financial risks on the other hand, are risks associated with the liberalization of foreign exchange and domestic financial markets, the domestic and overseas expansion of banks and their venturing to other segments of the financial service sector. Non-traditional banking risks include market risks, interest rate risk, liquidity risk, price risk and others.

On the other hand, Cade (1999) views solvency risk, liquidity risk, credit risk, interest rate risk, price risk and operating risk as the most important types of banking risks. This is because these types of risks arise from the business of banking (Uremadu, 2000).

to Contents, page 1

From the foregoing, one will be wondering where are the other types of risks such as systemic risk, market risk, legal risk, and so forth. But, in his (Cade's) argument all fit into the six risks broad types. For instance, he argued that "the umbrella termed 'market/position risk', which bridges some of the interest rate and price risks mentioned above, has found favour in dealing and regulatory circles as a 'territorial patch' and conceptual counterpoint to credit risk (e.g. in the dialogue on capital adequacy on emanating from Brussel and Basel). One is therefore attempted to adopt it as the lingua franca the objection is that it is simply too sweeping as title, that it oversimplifies and is likely to be misinterpreted by the general reader as embracing many things that it is does not (Cade, 1999).

Whereas, Robinson (2009) argues that it is generally accepted that there are three main categories of risk in the financial services industry. These three main categories include: credit risk, market risk and liquidity risk. It could be inferred that he summarized these risks around well known shades of financial risks.

According to Pyle (1999) risk, in this context, may be defined as "reductions in the firm value due to changes in the business environment". He went further to argue that the major sources of value loss are identified as market risk, credit risk, operational risk and performance risk.

Risk and reward

From what we have seen previously in this study, it is now apparent that banks and other major operators in the financial services industry operate in a high-risk business. In other words, financial institutions are highly geared financial risk takers. Also taking these risks responsibly can be said to be the business of the management of the banks.

Financial institutions that run on the principle of avoiding all risks will be stagnant and will not adequately service the legitimate credit needs of the society (Sanusi, 2007). A banker or an institution doing a banking business cannot just avoid taking risks because the business of banking involves trust level and the trust is relative or differs and cannot be fully guaranteed as it involves human elements whose hearts are prone to or filled with several devices to outwit one another in a competitive world. The Holy Bible, the Bible, established that "The heart of man is deceitful and desperately wicked: who can know trust it? Jeremiah 17:9 in King James (2007) the answer to this question is nobody".

Again, according to Furash (1994) in Azeez (2010) banks make money by taking risks and lose money by not managing risks effectively. He went further to argue that for banks to produce superior shareholder returns in current markets, they must take on higher levels of risks than in the past (Okafor, 1983 and Uremadu, 2000). Whereas, on the other hand, Sanusi (2007) says with proper risk management can lead to the bank running into murky waters. This is evident from so many bank's losses and in some cases even bankruptcy. Typical example of these issues can be traced back to as far back as 1980s. Merill Lyrich lost $377million in 1987 through trading mortgage-backed securities in an innovative form; Midland Bank also lost a reported £116million by guessing on wrong interest rates movements. In 1991, Bank of New England made a huge bad debt provisions, suffered a run on deposits and had to be supported by the government to the tune of $2billion. In 1992, Barclays Bank provided for bad debt of £2.5billion, which made it to declare its first loss in the history of its operations. Barving Bank, London's oldest merchant bank collapsed as a result of losses of $830million on speculative propretary position in Nikkei 225 stock index futures, the collapse of Lehman Brothers and the subsequent issues that led to the most recent banking crisis of 2007/2008-2009, which is still yet to fully recover from that singular scenario.

According to, Moyi (2010), banks over the past few years have taken excessive risks in their business activities in order to compete, more especially the Nigeria banks have also over indulged in this uncontrolled act leading to collapse of many of them as earlier cited elsewhere in this paper. He further argues that these issues ranged from giving loans to public office holders with their positions as collaterals, poor management styles, poor internal controls, insiders abuses, corporate governance issues, financing illicit contracts among others (see Azeez, 2010).

In the process of the provision of their service, that is, the day-to-day activities of the business, the banks assume various kinds of risks as seen earlier in this paper. It is also important to note that the risks that surround the bankers' principal activities i.e. those involving its balance sheet items and other basic business of banking are not all borne by banks. Banks should therefore endeavour to always adequately diversify the risks they carry in the normal business of banking.

Concept of risk management

At this juncture it is important that we introduce the concept of risk management proper in this discourse. Over the last decades we have seen a marked development in what has become known as risk management. Techniques and methods have been developed and redefined which allow for risks to be identified, their effects evaluated and the most efficient means of control discovered.

According to Dickson (1991) and Jobst (2007), risk management takes a broad view of the problems posed by risk. He went further to argue that it starts by asking basic question such as what risks is the organization exposed to. It moves down from there to evaluate the likely impact on the firm by looking at the severity and frequency of occurrence of the risk. Having identified the risk, and evaluated it, risk management techniques are then applied to decide how the risks that have been indentified can be mitigated. This is the long and short of a risk management concept or framework that business organizations like the banking firm should adopt in its operations to shore up its performance.

From the foregoing, we can establish that the concept of risk management is embedded in the following process (a) Risk Identification – risk management takes the view that the firm is exposed to risk in a variety of ways and any one such way may cause financial loss (see Dickson, 1999). Several steps are therefore taken using established risk management techniques to highlight all areas where the company is likely to suffer. (b) Risk Evaluation – this is the second stage in this management process of evaluating the impact of risk on the firm. Dickson (1993) states that often this evaluation is made in a quantitative manner, that is, without the use of quantitative analysis. In this case, a risk manager could, for example, study a flowchart (which is a statistical instrument anyway) and make certain quantitative evaluations as to the effects which specific events may have. This kind of qualitative risk evaluation is something which benefits from experience and those involved in risk management invariably falls back on their own experience of similar events or situations in measuring potential impact of risk. (c) Risk Control – Dickson (1993) identifies risk control to be the final stage of the risk management process, in which he describes it to be in two forms, which are physical and financial. At the end, it is the economic control of risk that is the main objective of the risk management department. After identifying and evaluating the risks faced by the firm, the final stage is the most important where all the information received will be used in order to mitigate (or control) as Dickson (1993) calls it the risks faced by the business. The techniques highlighted are on three tiers as follows:

Elimination: which simply refers to not undertaken the project that is associated with the risk, for instance, including another factory. Which could be eliminated by not undertaken the project.
Minimization: this is, in most cases, when the risk cannot be eliminated; we are therefore left with the option of minimizing its impact. Dickson (2010) further states that less prevention is primarily concerned with minimizing risk, or uncertainty of loss.
Transfer: this is another risk measure also identified by many writers (Azeez, 2010). This is a way of transferring the risk of loss to another party. A typical example is the insurance in which insurance premiums are paid to insurance companies to cover a particular thing say a building against fire, which in the event of fire the insurance company will cover all the expenses of the building another or repairing that particular building depending on the contract.

Risk management in the banking industry

Risk management is the process by which managers satisfy those needs by identifying key risks, obtaining consistent, understandable, operational risk measures, choosing which risks to reduce and which to increase and by what means and establishing procedures to monitor the resulting risk position (see Pyle, 1999). The foregoing treatise could be termed concept of risk management in a nutshell. Risk management has become a hot topic in the financial sector world over a period of years. However, the recent banking crises of late 2007 and early 2008 till date have made both regulators and institutions to rush into major efforts to upgrade their risk management systems, and focus the management attention towards appropriate process for due consideration of the trade-off between risk and return in financial management decisions or issues (Okafor, 1983 and Uremadu, 2000).

To explain the techniques used by the financial institutions, one must begin by explaining the risks within the banking industry and those that the institutions have chosen to manage. At this point, it is important to mention that not all risks that surround the industry are being managed. According to Santomero (1999) in Azeez (2010), have recognized that they should not engage in business in a manner that unnecessarily imposes risk upon them; nor should they absorb risks that can be efficiently transferred to other participants. In this case, he argues that banks should accept only the risks that are uniquely a part of bank's array of unique value-added services.

According to Oldfield and Santomero (1997) in Azeez (2010), it has been argued that risks facing all financial institutions can be segmented into three (3) separable types from a management perspective. These include:

i.) Risks that can be eliminated or avoided by simple business practices.
Ii.) Risks that can be transferred to other participants, and
iii.) Risks that must be actively managed at the firm level.

In the first case, the practice of risk avoidance involves actions to reduce the chances of losses from banking activities by eliminating risks that are superfluous to the institutions business purpose. These risk avoidance practices comprise of three types of actions. These actions include the standardization of processes and procedures to prevent inefficient or incorrect financial decisions. The second is the construction of portfolios that benefit from diversification across borrowers and that reduce the effects of anyone loss experience is another. Lastly, the implementation of contracts with management that require employees to be held accountable.

In the second case, it is the technique of risk transfer or reduce highly the inherent risks in their positions. Different markets exist for many of the risks borne by banks. A typical example of this can be easily transferred by rate sensitive products such as swaps and other derivatives (Azeez, 2010). It is now left to the Nigerian banking firms to take advantages and opportunities offered by these markets in a bid to utilize and or apply

to Contents, page 1

better risk management methods embedded in the prescriptions of corporate governance principles. It is high time the Nigerian banker re-orders his properties aright in the business of banking practice in the 21st century world.

SUMMARY OF FINDINGS AND RECOMMENDATIONS

At this juncture, we are going to summary our findings and discoveries from the paper and simultaneously state recommendations alongside them. These are as follows:

1. We discovered from the study that banks and other operators in the financial services industry engaged in high-risk businesses, that is, they are risk-takers. Consequently, it is here and recommended that they should try to moderate their risk taking positions, and always apply the three main techniques of risk performance and profitability.
2. We found that three main categories of risks inherent in the financial services industry or sector are; (i) credit risk (ii) market risk and (iii) liquidity risk. It is therefore advisable that Nigerian bankers be mindful of these risk categories, identify, diversify, manage or avoid them where possible to raise performance and profitability in the economy.
3. Evidence from the study also reveals that financial institutions do not completely avoid risks rather they manage risks towards growing returns and or profitability. This is encouraging as well as instructive to all operators in the industry that what is important is positively managing or avoiding risks but not completely avoiding them.
4. That Nigerian banks indulged (engaged in excessive risk-taking and never cared to apply risk management techniques over the years thereby leading to total collapse of majority of them. (From 72 banks to now 21 banks 3 just collapsed on 6/08/2011 most recently out of the 8 problem banks Central Bank of Nigeria (CBN) asked to recapitalize in 2009.
5. Study established causes of bank failure arising from poor risk management are: giving loans without proper appraisal and documentations, giving loans to public office holders with their positions as collaterals, poor management styles, poor internal control systems, insider abuses, corporate governance issues, financing illicit contracts and so forth. In this we make quick to recommend that banks should avoid or abstain from all improper activities during lending that often lead to credit defaults and or non performing loans, such as improper documentation/appraisals, poor management style, poor internal controls, insider abuses, “spinning”, illicit trade/contracts and avoid round tripping. Above all, they should instill practice of good corporate governance in themselves.
6. Risk management techniques, concepts and principles have been discovered over the years for proper bank management practices. It is therefore recommended that risk management technique should be properly employed in the business of banking to achieve direct results and or returns.
7. Banks carry various kinds of risks in their normal business of banking such as lending and borrowing and other financial services activities. It is therefore expected that banks should always adequately diversify the risks they carry in the normal business banking.
8. Finally, paper established that risk control process involves (a) risk elimination (b) risk minimization and (c) risk transfer. We recommend application of a mix of the three processes of risk management techniques by bankers and other operators in the financial services sector or financial institution in a bid to eliminate, minimize and transfer risks in lending and borrowing activities to raise their performance and growth of the national economy, in general.

CONCLUSION

In this paper, we have discussed the concept of risk management techniques and corporate governance and their applicability to business operations in Nigeria with a view to raising bank’s performance in particular and national productivity, in general. Using descriptive industry and economic analytical models we investigated and established that Nigerian banks engaged in excessive risk-taking; bank failures are caused by lack of application of risk management techniques by bankers and that banks carry various kinds of risks in their normal business of banking. We therefore conclude that if Nigerian banks fully apply risk management techniques and concepts embedded in the basic principles of corporate governance, banking sector performance will significantly improve thereby leading to growth of the domestic economy in the 21st century emerging economy to which Nigeria belongs.

REFERENCES

Azeez M (2010) “The Relevance of Risk Management in Business Operations in Nigeria”. An Unpublished M.Sc. Dissertation submitted to the department of accounting and economics, faculty of business studies, University of oxford, UK, Britain, in partial fulfillment for the requirement for the award of an M.Sc Degree in Accounting and Economics.

Cade E (1999) Managing Banking Risks. Cambridge, England: Gresham Books.

Cade E (2001) Managing Banking Risks: Reducing Uncertainty to Improve Bank Performance. Chicago, United States: Fitzroy Dearborn Publishers.

Dickson A (1991). Chartered Insurance Institute Study Text. London: BPP Publishing Limited.

Hicks HG, Gullet CR (1981). Management New York: McGraw-Hill Book Company.

Jobst AA (2007). “Operational Risk: The Sting is Still in the Tail but the Poison Depends on the Dose”. IMF Working paper WP/07/039. Retrieved from http://www.inf.org/external/pub/ft/wp/2007/wp 07239.pdf

King James Version (2007). Holy Bible, London: Worldwide Printing Company.
Okafor FO (1983). Investment Decision: Evaluation of Projects and Securities". London: CASSELL Ltd.
Philippatos GC (1970). Essentials of Financial Management: Test and cases. San Francisco: Holden – Day, Inc.
Pyle D (1999). Bank Risk Management: Theory and Practice. Norwell, USA: Kluwer Academic Publishers.
Robinson S (2009). Operation risk – CISI – TAQ Programme; Unit 10 Workbook. London: Chartered Institute for Securities and Investment.
Sanusi LS (2007). "Minimizing Risk Exposure in Nigerian Banks". Bulletin (Publication of Central Bank of Nigeria) 31: 3.
Sharpe FW, Alexander GJ (1990). Investments. Englewoods Cliff, New Jersey: Prentice – Hall, Inc.
Uremadu SO (2000) "Financial Management (Concepts, Analysis and Applications). Enugu: Precision Publishers Limited.
Uremadu SO (2009) "Bank Capital Structure, Liquidity and Profitability: Evidence from the Nigerian Banking System". Paper Presented at the 6th International Finance Conference (IfC6), 10-12 March, 2011, Harrimamet.

Prime Journals of Business Administration and Management (BAM)
ISSN: 2251-1261
Vol. 2(1), pp. 405-411, January 4th, 2012
http://www.primejournal.org/BAM
to Contents, page 1

Review

Corporate governance in Islamic financial institutions: An ethical perspective

Zulkifli Hasan

Senior Lecturer, Islamic Science University of Malaysia

Accepted 17th November, 2011

While significant concerns have been invoked on the material aspects of Islamic finance such as financial growth and products sophistication, it is nevertheless observed that equal emphasizes have not been given on the issue of ethics. In view of scarcity of literature on the subject and the essence of ethics in Islamic finance, this article aims at expanding the faith based moral horizon by advocating ethics as one of the foundational dimensions of corporate governance in Islamic Financial Institutions (IFIs). Unlike the western concept of corporate governance which is based on the western business morality that derived from "secular humanist", this article suggests that corporate governance in IFIs is founded on the epistemological aspect of Tawhid, Shari'ah and ethics.

Key words: Corporate Governance, Shari'ah, ethics and IFIs.

INTRODUCTION

Islam is neither simply a religion nor a mere ideological vision. It is a practical system of life and balance between human bodily requirements, spirit and reason. Islam is a comprehensive religion and it covers aqidah (belief), Shari'ah and akhlaq (ethics). In contemporary perspective, the ideal practice of Islamic finance tends to provide evidence about the comprehensiveness of Islam by invoking these three core elements.

Contrary to the ideal assumption that Islamic finance is about belief, Shari'ah and ethics, it is observed nevertheless that in actual practice, Islamic finance is more anxious on the legal and mechanistic aspect of Shari'ah compliant. At this point, Balz, (2010: 250) views that Islamic finance is now experiencing a "formalist deadlock" where the industry is more concerned with formal adherence to Islamic law instead of promoting Islamic ethical values. This is affirmed by El Gamal, (2006) when he severely criticized the practice of Islamic finance particularly by highlighting the issue of Shari'ah arbitrage. Significant criticisms by numerous scholars about the current practice of Islamic finance have led to series of questions as to the distinctiveness of Islamic finance with its conventional counterparts. Chapra, (2010) and Siddiqi (2007) for instance view that the practice of Islamic finance seems unable to attain its authenticity and share many common similarities with conventional finance. As a result, Islamic finance industry is also experiencing the impact of recent financial crisis such as in the case of closure of Ihlas Finance House in Turkey, the Islamic Bank of South Africa and Islamic Investment Companies of Egypt. These corporate failures raise an issue on the importance of ethics as the core element of Islamic finance.

There are numerous observations about major causes of corporate difficulties as experienced by several IFIs and one of them is weak of corporate governance (The OECD, (2004: 11) defines corporate governance as "a set of relationship between company's management, its board, its shareholders and other stakeholders, thereby plays essential function to provide the structure through which the company goals are set and the means of attaining those objectives and monitoring performance are determined". The IFSB, (2006: 33) defines corporate governance "as a set of relationships between a company's management, its BOD, its shareholders and other stakeholders which provides the structure through which the objectives of the company are set; and the means of attaining those objectives and monitoring performance are determined"). The analysis of such corporate failures in the existing literature nevertheless is

to Contents, page 1

found to heavily emphasize on the aspect of regulatory failure, management failure and control failure in corporate governance structure. It is observed that there is lack of discourse on the issue of ethics or the ethical failure aspects in IFIs that lead to such corporate difficulties. In view of scarcity literature on this subject, this article aims at providing general overview and basic understanding on the ethical perspective of corporate governance in IFIs. This article adds the previous literature on ethics and corporate governance by highlighting the distinctiveness of Islamic ethical principles and possible mechanism to institutionalize it. The remainder of this article is structured as follows. The next section reviews western and Islamic literature on ethics and corporate governance and section 3 specifically explains conceptual framework of Islamic ethics and its underlying principles. Section 4 attempts to highlight possible approach to promote and implement ethics by advocating the institutionalization of ethics and finally, section 5 concludes the discussion.

Ethics in western and Islamic theory of corporate governance

Becht and Barca, (2001) provide a literature review of a number of corporate governance models as possible solutions to solving the collective action problem among shareholders such as takeover model, block holder model, board models, executive compensation models and multi-constituency models. Another examination of the existing corporate governance theory can be found in Lewis, (1999: 33-66) where he examines the Anglo-Saxon model, the Germanic model, the Japanese model, the Latinic model, the Confucian model and the Islamic model. Generally, all of these corporate governance theories are either developed on the basis of agency theory or stakeholder theory orientation.

The Anglo-Saxon model of corporate governance which is also known as market-based systems or shareholder-value system or principle-agent model is considered as the most dominant academic view. Miller (2004:2) views that shareholders value orientation concerns on the sovereignty of individual where sole consideration is given to shareholders. Stakeholders' value model on the other hand focuses on a relationship-based model that emphasizes on maximizing the interests of a broader group of stakeholders (Adams, 2003: 4). Both these Western theories of corporate governance tend to focus on the mechanism of resolving the agency problems. These theories nevertheless have failed to certain extent to take into account the element of ethics as an essential component of corporate governance. Only after the incident of significant corporate failures and financial scandals due to lack of ethical consideration, there were suggestions to integrate ethics into corporate governance framework such as Drennan, (2004), Cladwell and Karri, (2005), Arjoon, (2005) and Sullivan and Shkolnikov, (2007) (It is reported that the implementation of business ethics provides positive effects on firms' financial performance (Spiller, 2002: 150)).

The notion of integrating ethics as part of corporate governance system then raises an issue of philosophical foundation of ethics in conventional literature. Basically, the ethical dimension in western theory is built on the basis of utilitarianism, relativism and universalism (Beekun, 1996). The ethical principles that extracted from these theories are based on philosophical ethics which is constructed from social interaction. All of the universal ethical principles which are applicable to corporate governance such as accountability, transparency, fairness and responsibilities are socially constructed through human reason and experiences.

On the other hand, Islamic model of corporate governance advocates comprehensive approach by emphasizing the elements of ethics as propounded in al Quran and al Sunnah. Unlike ethics from western theory perspective, the Islamic ethical principles are divine and religious construct. Wilson, (2002: 53) states that the Islamic ethics as being enduring and based on holy revelation while the ethics in western theory derived from social values are more transitory in nature. Al Quran and al Sunnah provide guidelines and principles of ethics that can be universally applied including in the matter of corporate governance. The distinctiveness of corporate governance theories from Islamic and Western perspectives and its relationship with ethics can be summarized as follows:

Table 1 illustrates the diversity of Islamic model of corporate governance with its conventional counterparts in six main areas i.e. the episteme, right and interest, corporate goal, nature of management, management board and nature of business. Unlike the shareholder value system and the stakeholder value orientation, the Islamic model of corporate governance puts aqidah, Shari'ah and ethics as the main component of its framework (This is in line with the IFSB, (2006: 33) where it explains corporate governance in the context of IFIs to encompass "a set of organizational arrangements whereby the actions of the management of IIFS are aligned, as far as possible, with the interests of its stakeholders; provision of proper incentives for the organs of governance such as the BOD, *Shari'ah* Board and management to pursue objectives that are in the interests of the stakeholders and facilitate effective monitoring, thereby encouraging IFIs to use resources more efficiently; and compliance with Islamic *Shari'ah* rules and principles"). At this point, Chapra, (1992), views that the underlying Islamic ethical principle acts as a moral filter for socio economic justice. In this regard, a set of values and ethical principles as defined by divine revelation through al Quran and al Sunnah provides clear guidelines as to the ethical consideration that relevant to corporate governance particularly in setting the standard code of behavior of all stakeholders and guiding the daily and business activities of the firm.

Table 1: The Distinctiveness of Corporate Governance Theories from Islamic and Western Perspectives

Aspects	Shareholder Model	Stakeholder Model	Islamic Model
Episteme	Rationalism and Rationality Ethics is social construct	Rationalism and Rationality Ethics is social construct	Faith-based rationalism: Aqidah, Shari'ah and Akhlaq Ethics is holy revelation
Rights and Interest	To protect the interest and rights of the shareholders.	The right of community in relation of the corporation.	To protect the interest and rights of all stakeholders but subject to the rules of Shari'ah and ethics.
Corporate goal	Shareholders profit	Stakeholders and Social welfare	Maqasid Shari'ah (*Maqasid Shari'ah* means protection of the welfare of the people, which lies in safeguarding their faith, life, intellect, posterity and wealth (Al-Ghazali, 1937: 139-140)): Ethics is part and parcel of maqasid Shari'ah
Nature of Management	Management dominated	Controlling shareholder dominated	Concept of Khalifah (vicegerency), Shura (consultation) and subject to Shari'ah and ethics
Management Boards	One-tier board No specific committee on ethical issues	Two-tier boards; Executive and supervisory Board/ No specific committee on ethical issues	Shari'ah board or any institution that responsible on the Shari'ah and ethical issues.
Nature of Business	No specific restriction on any kind of business	less prioritized No specific restriction on any kind of business	Only Shari'ah and ethical permissible activities are allowed.

Ethical dimension of corporate governance from Islamic perspective

Conceptual framework of ethics

The word ethics is derived from the Greek word ethos, which means character or custom (Solomon, 1984:3). It represents a wide meaning of character, behavior or code of conducts. In Islam, the word ethic is synonym with the term adab and khuluq (Siddiqui, 1997: 423). These two terms denote good behavior or a standard of conduct to be observed in social interactions (Saedon and Kamal, 1992: 51-62) or the set of moral principles that distinguish right and wrong (Beekun, 1996: 2). In the holy al-Quran the term khuluq can be found in Surah al-Qalam verse 4 as Allah says: "And surely you (Prophet Muhammad) have the best form of morals," and in surah al-Shu'ara verse 137: "There is no other than khuluq of the ancient". Apart from these, the Prophetic hadith had also made reference to ethics and morality where Aishah reported that that "the Khuluq (Morals) of the Prophet was based upon the Qur'an" and the Prophet says that "I have come to complete the code of moral conduct" (Muslim).

In deconstructing the Islamic ethical principles within the realm of economic, Naqvi, (1981: 45-57) advocates four important axioms that specifically reflect its relevancy in determining the rules of economic behavior in a society. The axioms of unity, equilibrium, free will and responsibility are the basis for deriving a set of ethical system and principles that would be appropriate to nurture and guide the economic behavior from Islamic point of view (The concept of unity refers to vertical dimension of Islam whereby man's life on earth in its entirety relates eternally to God (Naqvi, 1981: 48). While unity depicts the vertical dimension, equilibrium denotes the horizontal dimension of Islam by which it is a binding moral commitment of every individual, institution, corporation or any kind of entities to uphold a delicate balance in all aspects of lives (Naqvi, 1981: 51). The axiom of free will then propagates the concept of natural freedom within certain limitation whereby it emphasizes on the element of balancing between the 'individual freedom' and 'collective freedom' (Naqvi, 1981: 52). Finally, the concept of '*amanah'* or responsibility complements the Islamic ethical axioms in which the natural freedom that derived from the free will axiom must be exercised with full responsibility as a vicegerent and trustee of God (Naqvi, 1981: 54)). These divine formulated axioms provide very useful guidelines in identifying and recognizing legitimate ethical principles in economic.

Another construct of ethics to legitimize the ideal Islamic economic behavior refers to the principle of adl (justice), amanah (trust) and ihsan (benevolence). Based on the ethical axioms of unity, equilibrium, free will and responsibility, Islamic ethics must at least have three important characteristics namely the criterion of adl (justice), amanah (trust) and ihsan (benevolence) (Beekun and Badawi, 2005: 134-135). The first feature of ethics in Islam requires all individual to behave justly to all (Allah says in *al Quran "Allah commands justice, the doing of good and liberality to kith and kin, and HE forbids all shameful deeds and injustice and rebellion: He instructs you, that ye may receive admonition"* (*Al Quran*, 16: 90)). The managers for instance shall treat equally

Table 2: Islamic Ethical Principles

Ethical Principles	Sources
Prohibition of Riba (Interest)	"O you who have attained faith! Remain conscious of God, and give up all outstanding gains from Usury, if you are (truly) believers" (Al Quran, 2: 278, 2: 275, 3: 140, 4: 161 and 30: 39).
Prohibition of Maysir (Gambling)	"O you who believe! Intoxicants and gambling, (dedication of) stones and (divination by) arrows, are an abomination of Satans handwork: Abstain from such (abomination), that you may prosper. Satans plan is (but) to excite enmity and hatred between you with intoxicants and gambling, and hinder you from the remembrance of Allah, and from prayer: Will you not then abstain?" (Al Quran, 5: 90-91).
Prohibition of Gharar (Uncertainties)	Hadith narrated by Muslim, Ahmad, 'Abu Dawud, Al Tirmidhi, Al Nasa'i, Al Darami and Ibn Majah on the authority of Abu Hurayra where the Prophet prohibited the pebble sale and the gharar sale.
Good Character and Behavior (Islam insists on obligation to observe good character and behavior as the Prophet himself is sent to us for purpose of perfecting the best of character.)	Abd Allah ibn 'Amr said, the Prophet used to say: "The best of you are those who have the most excellent morals" (Bukhari, 61: 23). Hadith narrated by Abu Hurairah, the Messenger of Allah said: "The most perfect of the believers in faith is the best of them in moral excellence, and the best of you are the kindest of you to their wives" (Al Tirmidzi, 10: 11). Muadh Ibn Jabal reported that the Prophet said: Fear Allah wheresoever you may be, follow up an evil deed by a good one which will wipe (the former) out and behave good-naturedly to people" (Al Nawawi, 2001: 35).
Generosity in Doing Business (Generosity in doing business refers to willingness and kindness in giving away either rights or property or times for the benefit of others)	Uthman bin Affan reported that the Prophet said: "Allah will admit to the Paradise a man who is lenient as a seller and a buyer." (Ibn Majah, 3: 2202) Jabir bin Abdullah reported that Allah's Messenger said: "May Allah have mercy on the bondsman who is kind when he sells, kind when he buys and lenient when he demands (his debt)" (Ibn Majah, 3: 2202)

the employees without discrimination. The concept of amanah then further characterizes Islamic ethics by considering individual as a vicegerent of God and He is accountable to Him (In *al Quran*, (8: 27) Allah says "*Ye that believe! Betray not that trust of Allah and the Messenger, nor misappropriate knowingly things entrusted to you")* in which requires him to be responsible in whatever he does. Finally, the concept of Ihsan represents the core and most important element of Islamic ethics. Unlike justice which is mandatory, Ihsan denotes what is above and beyond mandatory (Al Qurtubi, 1966) (*Adl* refers to the person's inner intentions and feelings that should be consistent with the declared words and actions, while *Ihsan* goes beyond that where it requires words, actions and intention of certain good deeds sincerely realizing he is accountable to Allah (Al-Qurtubi (1966, 10: 165, as cited in Beekun and Badawi, 2005: 134)). In this regard, Ihsan requires extra caution, effort and good intention where the individual performs good deeds with the realization that Allah is watching him at all times (In *hadith* narrated by Umar, the Prophet explained *ihsan* as the act of worshipping Allah as though you are seeing Him, and while you see Him not yet truly He sees you (Al Nawawi, 2001)). The criterion of ihsan then expects all stakeholders in IFIs regardless of shareholders, managers, board of directors (BOD) and employees to observe the set of Islamic ethical principles which is divinely revealed and clearly stipulated in al Quran and al Sunnah.

Ethical principles

In discussing ethics in the context of corporate governance in IFIs, this article highlights several Islamic ethical principles that relevant to key stakeholders and these include prohibition of riba, maysir and gharar, to observe good behavior and conduct with candor, courtesy and fairness, to acquire knowledge, diligence and competence, to uphold interest of all stakeholders, fair competition, transparency, confidentiality and fair price and wages. All of these ethical principles are extracted and derived from the Islamic ethical axioms of unity, equilibrium, free will and responsibility as well as the criterion of adl, amanah and ihsan. It is worth noting that this list is non-exhaustive. Table 2

The command to avoid riba, maysir and gharar is considered as main characteristics of Islamic ethical principles. The prohibition of riba, maysir and gharar is clearly mentioned in al Quran and al Sunnah. Unlike the western ethical code which is based on social interaction in which riba, maysir and gharar are tolerable, Islam vividly declares these three elements to be unethical and can not be compromised at all times. The remainder of the underlying Islamic ethical principles however may share some common similarities with the western ethical code.

INSTITUTIONALIZATION OF ETHICS IN IFIS

Institutionalization of ethics is one of the best approaches to promote and implement the Islamic ethical principles

Table 3 continues: Islamic Ethical Principles

Ethical Principles	Sources
Diligence and Competence (Islam also concerns on the quality of works where individual is required to improve their working standards to the utmost and it should be maintained constantly so as to avoid any occurrence of negligence which could harm the interest of others)	The Prophet said: "Allah loves to see one's job done at the level of itqan or wisdom" (Cited in Al Dimasqi, 2006: 385-388). Saidatina Aishah reported that the Rasulullah said: "The deeds most loved by Allah (are those) done regularly, even if they are small" (Bukhari and Muslim).
To Uphold Interest of All Stakeholders	Ibnu Umar reported that the Messenger of Allah said: "A Muslim is the brother of a Muslim; he does him no injustice, nor does he leave him alone (to be the victim of another's injustice); and whoever does the needful for his brother, Allah does the needful for him; and whoever removes the distress of a Muslim, Allah removes from him a distress out of the distresses of the day of resurrection; and whoever covers (the fault of) a Muslim, Allah will cover his sins on the day of resurrection"(Bukhari, 46: 3) Anas reported that the Prophet said: "Help thy brother whether he is the doer of wrong or wrong is done to him."They (his companions) said, O Messenger of Allāh! We can help a man to whom wrong is done, but how could we help him when he is the doer of wrong? He said: "Take hold of his hands from doing wrong." (Bukhari, 46: 4).
Fair Competition (It is universally affirmed that unfair competition is an act that against human right, dictum and principle of morality and ethic. The basic premise of ethics on competition is to promote fair trade, healthy competition and ultimately consumer welfare in the market)	The Messenger of Allah said: "Do not hate one another and do not be jealous of one another and do not boycott one another, and be servants of Allah (as) brethren; and it is not lawful for a Muslim that he should sever his relations with his brother for more than three days" (Bukhari, 78: 57). The Prophet declared that: He who monopolizes is not but a wrongdoer" (Al-Tirmidhi, 6: 23).
Transparency (Transparency is of utmost importance as *al Quran* specifically forbids concealing of evidence)	"And if you are traveling and cannot find a scribe, then there be mortgage taken.….and do not conceal not evidence for whoever hides it, surely his heart is tainted with sin and Allah is knower of what yo dou do." (Al-Quran, 2: 283). Abdallah reported that the Prophet said: "Truthfulness leads to righteousness, and righteousness leads to Paradise. And a man keeps on telling the truth until he becomes a truthful person. Falsehood leads to Al−Fajur (i.e. wickedness, evil−doing), and Al−Fajur (wickedness) leads to the (Hell) Fire, and a man may keep on telling lies till he is written before Allah, a liar" (Bukhari, 8: 116).
Confidentiality (Every individual has also duty to keep secret of all communications which is classified as private and confidential unless it is against the public interest and justice)	Hadith reported by Abu Hurairah, the Prophet said: "Whosoever relieves from a believer some grief pertaining to this world, Allah will relieve from him some grief pertaining to the Hereafter. Whosoever alleviates the difficulties of a needy person who cannot pay his debt, Allah will alleviate his difficulties in both this world and the Hereafter. Whosoever conceals the faults of a Muslim, Allah will conceal his faults in this world and the Hereafter. Allah will aid a servant (of His) so long as the servant aids his brother. Whosoever follows a path to seek knowledge therein, Allah will make easy for him a path to Paradise. No people gather together in one of the houses of Allah, reciting the Book of Allah and studying it among themselves, except that tranquility descends upon them, mercy covers them, the angels surround them, and Allah makes mention of them amongst those who are in His presence. Whosoever is slowed down by his deeds will not be hastened forward by his lineage" (Cited in Al Asin, 1970: 14-15).
Fair Wages and Price (Any practice of excessive price is against the Islamic ethical principle. Fair policy such as fair remuneration and equal employment opportunities is actually part of ethical consideration)	"Withhold not things justly due to others" (Al Quran, 26: 18). Hakim bin Hizam reported that the Prophet Muhammad said: "The seller and the buyer have the right to keep or return the goods so long as they have not parted or till they part; and if both the parties spoke the truth and described the defects and qualities (of the goods), then they would be blessed in their transaction and if they told lies or hide something, then the blessings of their transaction would be lost " (Bukhari, 3: 293).

to Contents, page 1

Table 4: Agent of Ethics in IFIs

Agent of Ethics	Functional Roles
Regulatory Authority	To set regulatory framework for sound and proper code of ethics - Code of ethics on corporate governance for general usage
Supervisory Authority	To supervise, monitor and enforce the implementation of code of ethics - Enforcement of the code of ethics
Shareholders	To ensure that all investments and business activities are Shari'ah and ethically permissible - Incentive for ethical achievement
Shari'ah Board	To ensure Shari'ah and ethical compliance - Assist the BOD to come out with Code of Ethics - Emphasize on ethics in the process of issuing Shari'ah rulings
BOD	To set the IFIs direction and policies on ethics - Code of Ethics for internal usage - Ethics as a basis of decision making
Management	To implement set of ethical policies set by the BOD - Organizing ethics training - Module for ethics programme - Enforcement of ethics
Employees	To practice and comply with the code of ethics - Ethics as a culture

as highlighted above in any organization. Basically, the process of institutionalization of ethics requires a formal initiative to guide key stakeholders in the corporation to implement and promote ethics. Such process is very important in order to control the problem of ethical issues in the corporations (Vitell and Hidago, 2006). The existing practice shows that institutionalization of ethics in corporation can be in the form of establishing permanent board-level committee that responsible to set the policy on ethics (Certain companies initiate the establishment of ethical committee known as "Defalcation Committee" to resolve cases related with ethical issues. This committee is an avenue within the internal structure of corporate governance in the firm to resolve disputes and cases involving violation of code of ethics (Othman and Abd Rahman, 2009: 380)), issuance of code of ethics, organizing ethics training, reinforcing the employee's organizational commitment, and encouraging an ethically-oriented organizational culture (Sim, 1991). All of these actions would be able to create awareness about ethics and at the same time to promote the implementation of ethics as part of corporate governance framework.

Any action and effort to institutionalize ethics adheres most to its key players within the corporate governance structure of the organization. This raises an issue as to the need for specific agent for such purpose. At this point, several key participants of corporate governance either external such as regulatory and supervisory authorities or internal as in the case of BOD, shareholders, managers, employees and Shari'ah board are considered as agents of ethics. Their responsibilities to promote, to implement, to practice and to enforce ethics are summarized in table 3.

The regulatory authorities play a key role in promulgating a set of law or code of ethics on corporate governance. To complement this function, the supervisory authorities have duty to supervise and monitor the implementation of this code of ethics effectiveness of corporate governance system and to check its. Shareholders have responsibilities to ensure that all business transactions and investment activities are conducted in ethical way. The BOD has responsibility to specify the code of conduct and standard of appropriate behavior for internal usage. Unlike the BOD, the management has fiduciary duty to implement the ethical policies and strategies set by the BOD while the employees, to practice and observe every aspect of ethics as stipulated in the code of ethics.

The most essential agent of ethics in IFIs is Shari'ah board. Basically, the functions of Shari'ah board are two-folds i.e. advisory and supervisory and these include advising IFIs in its operation, to analyze and evaluate Shari'ah and ethical aspects of any banking and financing activities and to monitor and supervise the extent of Shari'ah compliance. Considering to their expertise and knowledge on Shari'ah and the state of its independence, Shari'ah board shall play an active role to promote Islamic ethics and values within the organization. The existence of Shari'ah board within the internal corporate governance structure shall be the advantage for IFIs to further promote the implementation of ethics in daily business activities, decision making process, management style, financial products and services and etc.

CONCLUSION

In spite of its similarities and some common principles with the western theory of corporate governance, Islam adds additional value to the existing governance framework whereby it emphasizes on the element of faith, Shari'ah and ethics. The foregoing discussion on the ethical dimension in Islam further validates the need for inserting ethics as part of corporate governance framework in IFIs. Islam highly insists stakeholders of IFIs particularly shareholders, BOD, managers, employees and Shari'ah board to preserve standard of conducts, to observe good ethics, conscience and piousness. Since this article is theoretical in nature, further study is really needed in order to explore the extent of ethical practice in IFIs. This triggers the call for future research to complement the existing theoretical studies on ethics with appropriate empirical data and information.

REFERENCES

Al-Asqalani, Ahmad Ibn, Ali, Ibn Hajar. (2000). "Fathu al-Bari Syarh Sahih Bukhari". Volume 10. Damsyik: Dar al-Faiha'.

Al-Asin, Musa, Syahin. (1970). "Fath al-Mun'im, Syarh Sahih Muslim". Volume 10. Cairo: Maktabah al-Jamiat al-Azhariyah.

Al-Ghazali, Abu Hamid Muhammad. (1937). Al-Mustashfa. Vol.1. Cairo: Al-Maktabah Al-Tijariyyah Al-Kubra.

Al-Nawawi, Yahya ibn Sharaf. (2001). Text, Translation and Notes: Imam Nawawi's Collection of Forty Hadith. Kuala Lumpur: Islamic Book Trust.

Arjoon, Surendra. (2005). Corporate Governance: An Ethical Perspective. Journal of Business Ethics, 61 (4), 343-352.

Becht, M and Barca, F. (2001). The Control of Corporate Europe. Oxford: Oxford University Press.

Beekun, Rafeek. I. and Badawi, Jamal. A. (2005). Balancing Ethical Responsibility Among Multiple Organizational Stakeholder. Journal of Business Ethics, 60, 131-135.

Beekun, Rafik, I. (1996). Islamic Business Ethics. Herndon, Virginia: IIIT.

Caldwell, C. and Karri, R. (2005). Organizational Governance and Ethical System: A Covenantal Approach to Building Trust. Journal of Business Ethics, 58, 249-255.

Chapra, Muhammad, Umer. (1992). Islam and the Economic Challenge. Herndon, Virginia: IIIT.

Chapra, Muhammad, Umer. (2010). Innovation and Authenticity in Islamic Finance in Ali, Syed, Nazim, Islamic Finance, Innovation and Authenticity, (pp. 31-64). Cambridge, Massachusetts: Harvard University.

Drennan, L. (2004). Ethics, Governance and Risk Management: Lesson From Mirror Group Newsarticles and Barings. Journal of Business Ethics, 52 (3), 257-266.

El-Gamal, M. (2006). Islamic Finance, Law, Economic and Practice. New York: Cambridge University Press.

IFSB. (2006a). Guiding Principles on Corporate Governance for Institutions Offering Only Islamic Financial Services (Excluding Islamic Insurance (Takaful) Institutions and Islamic Mutual Funds. Kuala Lumpur: IFSB.

Lewis, M.K. (1999). "Corporate Governance and Corporate Financing in Different Cultures", in Zeljko Sevic (Ed.), Banking Reform in South East European Transitional Economies (pp. 33-36). London: University of Greenwich Business School, Balkan Center for Public Policy and Related Studies Humanities Research Centre.

Naqvi, S. N. H. (1981). Ethics and Economics, an Islamic Synthesis. Leicester: The Islamic Foundation.

OECD. (2004). OECD Principles of Corporate Governance (Revised). Paris: OECD. Available at: http://www.oecd.org/dataoecd/32/18/31557724.pdf. Access Date: 25th July 2010.

Othman, Zaleha and Abd Rahman, Rashidah. (2009). Recognizing Value Approach in Corporate Governance: Institutionalization of Ethics. International Review of Business Research, 5 (4), 374-387.

Saedon, Mahmud and Kamal, Nik Ahmad. (1992). Adab al-Qadhi in Daily Activities of a Qadhi. IIU Law Journal, 2 (1), 51.

Siddiqui, Ataullah. (1997). Ethics in Islam: Key Concepts and Contemporary Challenges. Journal of Moral Education, 26 (4), 423- 431.

Siddiqi, Muhammad, Nejatullah, (2006). What Went Wrong? Keynote Address at the Roundtable on Islamic Economics: Current State of Knowledge and Development of Discipline held at Jeddah, Saudi Arabia on 26-27 May 2004 under joint auspices of the Islamic Research and Training Institute, Jeddah; and the Arab Planning Institute, Kuwait. Available at: <http://www.siddiqi.com/mns/Keynote_May2004_Jeddah.html>, Access Date: 13th August 2010.

Sim, R. L. (1991). The Institutionalization of Organization Ethics. Journal of Business Ethics, 10 (7), 493-506.

Solomon, Robert. C. (1984). Morality and the Good Life. New York: McGraw-Hill.

Spiller, Rodger. (2000). Ethical Business and Investment: A Model for Business and Society. Journal of Business Ethics, 27 (1/2), 149-160.

Sullivan, J. and Shkolnikov, A. (2007). Business ethics: the key role of corporate
governance'. The Corporate Board, January / February.

Vitell, S. and Hildago, E. (2006). The Impact of Corporate Ethical Value and Enforcement of Ethical Codes on the Perceived Importance of Ethics In Business: A Comparison of US and Spanish Managers. Journal of Business Ethics, 64, 31-43.

Wilson, Rodney. (2002). Parallels between Islamic and Ethical Banking. Review of Islamic Economics, 11, 51-62.

to Contents, page 1
Prime Journal of Business Administration and Management (BAM)
ISSN: 2251-1261
Vol. 2(1), pp. 426-435, January 4th, 2012
www. primejournal. org/BAM

Full Length Research

Return on investment analysis: Applying a private sector approach to the public sector

LR Jones

Distinguished University Professor and George F. A. Wagner Professor of Public Management, Graduate School of Business and Public Policy, NPS, Monterey, CA USA 93942. E-mail: dukedmb@aol.com or lrjones@nps.edu

Accepted 19th December, 2011

In an environment of scarce resources and rising government deficits the public not only expects but demands greater accountability for the spending of public funds. This demand has created a trend in the public sector, not only in the United States, but worldwide, towards the importation of private sector business analysis practices to improve government accountability-oriented analysis. One example is increased emphasis on return on investment (ROI) analysis in public sector organizations. Development and application of ROI analysis is challenging in the public sector since most government organizations do not generate profit necessary for calculation of ROI in the manner in which it is done in the private sector. This paper addresses previous attempts at using ROI in the public sector, identifying whether these attempts properly used ROI and what prevented their ultimate success in terms of use value. This paper argues that properly designed and conducted ROI analysis, based on methods used successfully in the private sector, can better reveal how and for what goods and services public money is spent and a means for evaluating whether it was spent well in providing goods and services. The methodology developed in this study provides a means for comparing the value derived from investment and work performed by and for government using an approach to ROI based on private sector methods.

Keywords: Return on investment (RIO), public sector, private sector.

INTRODUCTION

Return on investment (ROI) is one of the key methods used to quantify the level of success achieved or achievable in a business endeavor. The concept of ROI is used throughout private industry not only to determine past results, but also to evaluate the current situation and as a decision making tool for the future. The advantages of ROI are clear in that it provides the flexibility to anticipate output changes in advance. This benefit results in the ability to not only preview the future in a real world sense, but also to modify the inputs to the numerator and denominator of the equation to model potential courses of action for the enterprise.

Although a useful concept, ROI does not easily transition for use in the public sector. Unlike private enterprise, the public sector has no “profit” or “total sales” to use in the equation (With respect to total sales, revenue isn’t directly relevant to ROI, but asset turnover ratios can be calculated). The increasing need for some method to quantify ROI in the public arena has led to multiple attempts from a diverse group of public enterprises with varying results (We may note this is the purpose of cost benefit analysis wherein, for service centers, through estimation of a shadow price a quasi profit may be estimated to produce a measure of ROI or EVA. This approach is not followed in this paper. On this methodology see Anthony E. Boardman, D. H. Greenberg, A. R. Vining, *Cost Benefit Analysis: Concepts and Practice,* (Upper Saddle River, NJ: Prentice Hall, 2001). The Australian government placed increased emphasis on what they termed the “value added” approach in an effort to determine the output they were receiving as a result of budgetary expenditures. The Royal New Zealand Navy desired a determination of ROI for the implementation of a retention bonus plan used to control the attrition problem that was being experienced with marine engineers. Both of these results were somewhat mixed with valuable lessons learned. The United States Postal Service (USPS) met with a greater level of success in their effort, due largely to the fact that they are run much more like a private enterprise. Although not seeking to be “profitable, ” the USPS does generate revenue which can be used in the numerator of the formula which when divided by the USPS asset base in the denominator results in a fairly conventional ROI. Finally, the US Navy Dental community effort was much more ambitious in that it attempted to convert non cash

Table 1: Selected Financial Data for PepsiCo: 2001 (in millions of dollars)

Earnings Before Interest and Taxes	$4, 181
Net Income (after Interest and Taxes)	$2, 662
Total Assets	$21, 695
Total Shareholder Equity	$8, 648

Source: Author, 2009

outputs into cash equivalents in order to closely adhere to the traditional ROI formula. The resulting Navy Dental ROI was dogged by the questionable accuracy of some inputs, but the overall approach remained fundamentally sound.

The previous efforts focused on in this study have helped define the need for a new method to determine ROI in public sector enterprises. The intent of this paper is to review the previous efforts and to develop a new approach for attaining this important goal.

Traditional versus notional ROI

Traditional ROI

ROI has traditionally been measured in the private sector to quantify an organization's past, present, and potential future performance. There are several methods by which an organization can determine its ROI. Most compare the net financial output of a company, or profit, to the financial input. One of the most common methods is to compute a percentage return on a company's assets. An organization can determine how efficiently it has used its assets by comparing a period's operating income to the total amount the company has invested in the assets that produced that income. ROI is traditionally calculated as follows (Ray H. Garrison and Eric W. Noreen, *Managerial Accounting* (San Francisco: McGraw-Hill Irwin, 2003), 542. It may be noted that private sector finance researchers sometimes refer to this method as "accounting ROI" (or Return on Assets -- ROA)):

$$\text{Return on investment} = \frac{\text{Net operating income}}{\text{Average operating assets}}$$

Net operating income is the difference between revenue and expenses, usually before taxes and interest. An average asset base is normally used since the amount of assets in use may have changed during the period of measurement. Regardless of the exact method of measurement, a higher return indicates a more proficient use of organizational assets and ultimately a higher return for its shareholders.

ROI calculations also may be used to determine the potential reward of a single investment decision or to assist in choosing between multiple investment options. For a single investment decision, forecasted streams of revenue are estimated and compared to the expected capital investment and operating costs. Under traditional formulations these are compared over the life of the proposed project and used to determine an internal rate of return (IRR), actually a forecasted ROI. The IRR is then compared to a firm's cost of capital for a single investment decision. It can also be compared to the IRR forecasted from other investment decisions in the case of multiple options to assist in choosing between them. With reasonable forecasting accuracy, this becomes an effective tool used in the private sector for deciding between capital venture decisions.

There are two methods frequently used to determine a corporation's ROI. Consider an investor deciding whether or not to make an investment in PepsiCo in 2002. One method for estimating PepsiCo's future performance is to look at its previous year's use of assets. This is commonly referred to as an organization's return on assets, or ROA. This was the method discussed in the previous section. Selected data for PepsiCo taken from 2001 are presented in table 1 (Richard et al. 2004).

To determine PepsiCo's ROA for 2001, their earnings before interest and taxes are taken from the income statement and must be divided by their total assets from the balance sheet. The result is then multiplied by one hundred to produce a percentage ROA. Thus, for PepsiCo in 2001;

$$\text{ROA}\,(\%) = \frac{\$4.2\text{B}}{\$21.7\text{B}} \times 100 = 19.3\%$$

This indicates that every dollar invested in assets at PepsiCo yielded 19. 3 cents of return in 2001 (The best numerator is free cash flow, with depreciation added back in).

Another, more relevant method to the investor would be to determine PepsiCo's return on equity, or ROE for 2001 (A significant difference between public and private entities is the method of financing. A private entity can be financed through both debt and equity, which leads to a difference between ROI and Return on Equity or ROE. The method of financing a public entity can be taken into account in choosing ROE over ROI, such as when a project is financed through both bonds and taxation, but this subtlety is not explored further in this paper). This explicitly gives the return on investor equity in PepsiCo. ROE is determined by dividing the net income (income after interest and taxes) by the corporation's total shareholder equity. The result is then multiplied by one hundred to produce a percentage ROE. Thus for PepsiCo in 2001;

$$\text{ROE}\,(\%) = \frac{\$2.7\text{B}}{\$8.6\text{B}} \times 100 = 31.4\%$$

This indicates that every dollar invested in PepsiCo by investors yielded 31. 4 cents of return to the shareholders in 2001.

to Contents, page 1

It is important to keep in mind that taken on its own the ROI is of limited value. In this example it would be wise to compare the ROA and ROE for PepsiCo to prior years or with other companies in the same business during the same year. This comparability is very helpful in determining if the ROI is superior, average, or mediocre. Careful evaluation of the inputs to the ROI formula can uncover what may be the root of the success or problem.

Public sector ROI calculations are considerably more problematic to utilize than in private industry. The traditional method of determining investment returns in the private sector is not directly compatible with many public sector organizations. Consider how a public sector organization would determine its financial output. Many public sector organizations do not produce revenues or generate profits as outputs. Therefore, their outputs are difficult to quantify in dollars. Instead, they provide a service or capability to the public. Oftentimes this service or capability is unique to the public sector and is not produced by the private sector. This increases the difficulty when trying to value these unique services or capabilities. For example, how much value is added to the respective service when another tank or fighter jet is produced? Certainly these costs are known, at least approximately. However, it is difficult to quantify their value added to the Army or Air Force. Placing dollar values on these items is complex since similar items are not valued in the private sector. The value added to the services from these items cannot be easily measured in dollars. This makes the use of traditional ROI criteria impossible (There is a way around this problem in some cases. For example, for a municipality it may be argued that there is a direct analog to private sector share price, that is. , the market value of the land within the jurisdiction's boundaries).

Some public sector organizations could be measured by the equivalent value of the service or capability provided in the private sector. For example, a comparison could be made between the United States Postal Service and United Parcel Service of America, Inc. Perhaps a cost comparison for compatible services between private and public sector organizations could be used to measure performance. However, many public sector companies do not have comparable organizations in the private sector. For example, consider the Department of Defense. The DoD provides defensive and offensive capability for the United States. This capability cannot be measured against the private sector due to the uniqueness of the services it provides. Therefore, to facilitate a ROI metric for many public sector organizations, a different approach needs to be used.

Notional ROI

ROI measurements are under exploration in the public sector for three primary reasons. First, as with private sector corporations under specified circumstances, there are always significantly more investment opportunities than public funds available (Despite the common view that this is always is the case for private sector firms, this assumption is not made in private finance. The assumption is that if an investment is wealth creating it should be made. This logic is even more relevant to the public sector, although as we know from our understanding of public budgeting, typically there are liquidity issues that prevent some investments that would produce positive returns from being made, especially in fiscally constrained jurisdictions). There is intense competitive pressure between organizations to continually prove their need for additional or even continued program funding. Deciding between these alternatives is oftentimes subjective in nature since objective data is not available. Realistically, some public programs will be funded regardless of their ROI. However, ROI measurements could provide one metric to objectively decide between investment alternatives in public programs. They could also be used by organizations to show their value added to the public, and consequently provide support for their continued funding. Second, increased public spending and rising budget deficits have considerably raised the public's concern for the way the public sector spends its money. There has been a notable increase in the required accountability of the public sector to the taxpayers. Evidence of this is the Government Performance and Results Act of 1993. The general purpose of this legislation is to establish metrics within the United States government to hold organizations accountable (Phillips and Phillips 2002). ROI measurements are one way this accountability requirement to taxpayers might be better met. Finally, there is a long-term trend importing selected successful business practices from the private sector and adapting them for use in government, for example. , from PPBS in the 1960s to just in time logistics in the 2000s. This is no surprise since many elected officials and public sector leaders have had previous careers in the private sector. Further, the private sector is viewed by much of the American public as more efficient than the public sector. The public appears to assume that unless a private organization produces a unique product or service that is in demand and meets customer preferences efficiently the firm will not survive. Despite examples to the contrary, there is the perception that private firms must perform efficiently to survive in competitive markets. While market pressures are not present in the public sector to the same degree, fiscal exigencies such as those present when the economy is in recession are likely to promote efforts to increase efficiency and cost effectiveness in the public sector. ROI measurement is a private sector financial measures that may be found useful in application in the public sector.

Past experience demonstrating the inability to apply ROI techniques to many organizations in the public sector indicates that to do so successfully require a different approach than used before. One such approach includes the use of cost effectiveness analysis to provide a useful framework with which to assign weights to the numerator variables. Boardman et al. (2001) addresses this issue as follows:

If the analyst is unable…to monetize the major benefit, then cost effectiveness analysis may be appropriate. Because not all of the impacts can be monetized, it is not possible to estimate net benefits. The analyst can, however, construct a ratio involving the quantitative, but non-monetized, benefit and total dollar costs (Boardman et al. , 2001).

This is the approach used in this paper, with the resulting formula producing a non-monetary output considered to be the notional return on investment, or NROI (The first chapter of Boardman et al. , (2001) explains why ratio measures are usually inappropriate. The authors argue that cost benefit analysis is clearly preferable to ratio analysis for use in the public sector. However, this paper argues that this is not always the case). In order for the NROI formula to be of credible value, weights must be assigned to each of the numerator variables. Weights are indicative of how the decision makers prefer to balance the impact of the attributes. This step is extremely important since the weight distribution has a tremendous impact on the output. Determination of weights can be an objective result of models and data analysis, a subjective result of discussion by the decision makers, or a combination of both. There are four common methods for determining weights: equal weighting, rank reciprocal, pair-wise comparison, and direct assessment.

The equal weighting method simply assigns equivalent weights to all of the variables. The rank reciprocal method has four steps. First, each variable is ranked in order of relative importance. Next the reciprocal of the ranks is taken (1/1, 1/2, 1/3, etc). The resulting fractions are then added together using a common denominator to create a new base (60/60 + 30/60 + 20/60 + 15/60 + 12/60 + 10/60 = 147/60). Finally, the original reciprocals for each variable are divided by the new base (147/60) with the resulting distribution being used for weighting. The equal weighting and rank reciprocal methods generally do not provide a high enough level of subjective scrutiny to be of value in a detailed project. With the pair-wise comparison method, the decision makers are provided a specific number of points to be distributed as they see fit between the variables. After discussion, each variable is assigned a numerical value. The sum of the values is then used as the denominator for the variable weighting, with the numerator being the assigned numerical value. Like the previous two methods; pair-wise also fails to provide enough ability to fine-tune the weighting distribution for a detailed project. The direct assessment method uses deductive reasoning to determine and assign weights to each variable. Although this method is purely subjective, it is less random and can easily be modified as necessary. The subjective nature of direct assessment can be alleviated to some extent by using a number of technical experts to develop the weighting values to be used in the formula.

Once weightings have been assigned, the non-monetized value of the numerator variables can be determined. After adding the variables together, the resulting numerator value is divided by the asset base to provide an NROI output. The validity of this NROI on its own is minimal. Trend analysis is required, using subsequent alterations to the numerator variables for the first ship in the class as it is completed and compared with independent data from the second ship in the class as it progresses. Essentially, the first NROI developed sets a baseline that is used to compare with subsequent outputs for the same project. The trend data from the first project can then be analyzed to determine if priorities can be adjusted for the second project in order to improve the output. Additionally, comparisons can be made between projects provided the category modifiers are the same and the scope of the project is similar.

While it is important to provide weights to the categories used in the numerator of the equation, it is even more important if at all possible to use monetary values for the NROI formula. In *The Bottomline on ROI*, while discussing the determination of ROI in the public sector, Patricia Phillips states that “converting data to monetary benefits is critical…the process is challenging, particularly with soft data, but can be methodically accomplished” (Phillips and Phillips 2002). For the formula to be most applicable for the purpose of comparing the past, present, and future NROI of a number of projects, it was determined that the effort must be expended to convert the data to monetary values. This would also serve to provide an NROI as near to traditional as possible for a public sector organization.

Previous efforts at developing ROI in the public sector

Australia

In an attempt to justify government acquisitions, Australia’s government continues to focus on improving its own ability to develop and implement ROI criteria. In a society vigorously competing for scarce public resources, receiving the best value for money spent has become central to government policy. The Australian Commonwealth demands this accountability. In response, Australia’s government has taken strides to emphasize the development of ROI criteria to make acquisition decisions throughout its governmental departments. However, the research supporting this study was unable to find a specific example where ROI criteria were successfully developed and implemented by a governmental organization. Dr. Allen Hawke, Australia’s previous Secretary of Defense, acknowledges the public’s frustration with their lack of success to date. According to Dr Hawke’s address in February of 2000, “there is a widespread dissatisfaction with Defense’s

to Contents, page 1

Performance (regarding Australia's Defense Organization use of funds). In essence we have a credibility problem" (Hawke 2000).

Australia's Department of Finance and Administration (ADOFA) is responsible for providing direction to Australia's ministries in making procurement decisions. Instead of simply choosing the lowest cost alternative, ADOFA emphasizes the "achievement of value for money" (Australia Department of Finance and Administration, 2003). Among other things, this method weighs the ability of the alternatives to meet the stated objectives, the reliability and reputation of the contractor, and the whole of life costs instead of just the initial procurement cost. Instead of providing structured guidance to determine ROI, ADOFA provides a substantial list of things to consider and leaves it to the particular agency to identify and weigh those things that apply. Due to the unique benefits of each procurement decision, this general approach may be appropriate. However, recent comments from Australia's Defense Procurement Review indicate a lack of success thus far within the acquisition community. The review concludes with the following comment;

Our review of the acquisition process has led us to conclude that there is no single cause of the failures that have become apparent in the development of capability and the acquisition and support of defense equipment (Australia Dept. of Defense, 2003: 47).

The Australian government does acknowledge the need to consider ROI when making procurement decisions. However, by merely emphasizing value for money in broad terms, they are not actually implementing quantitative ROI criteria within their government. They do highlight the need to consider many important factors other than costs for procurement decisions such as quality and contractor performance. Yet, they do not provide a universal method for considering the weighting of these factors so that decisions can be consistently made the same way across the different ministries. Perhaps this inconsistency is one of the reasons for their continued lack of success within the Department of Defence acquisition community.

New Zealand

The following ROI case concerning the Royal New Zealand Navy (RNZN) is summarized from a case study authored by Beryl Ann Oldham, Paul Toulson, Brenda Sayers, and Graham Hart. The RNZN had encountered considerable difficulty retaining their marine engineers (ME) in the mid 1990s due to high attrition rates. The ME community is responsible for many of the complex systems aboard the RNZN's fleet ships including operation and maintenance of diesel engines, gas turbines, electrical generators, and air conditioning and refrigeration plants (Royal New Zealand Navy, 2004). The attrition problem was so significant that the ability of the RNZN to maintain an acceptable operations tempo was threatened. Several suggestions were made in an effort to reduce the ME attrition rate. These measures included improved ME career management initiatives, better management of leave and maintenance periods, improved pay, compensation time for working weekends, and the more controversial Marine Engineer Retention Bonus Scheme (MERBS). It was believed that implementation of an immediate retention bonus was imperative to control the attrition problem in the short run since the other proposed initiatives would be slower to take effect (Oldham, et al. , 2002).

The MERBS was an expensive human resource endeavor for the RNZN. MERBS costs included both administrative program set up costs and the retention payments to personnel themselves. These overall costs were estimated at almost five million Australian dollars (144). However, the MERBS was considered a successful initiative since it did reduce attrition rate for the MEs to an acceptable level. Unfortunately, it was difficult to determine just how successful the MERBS was, especially considering that other retention initiatives were occurring simultaneously. Consequently, an ROI study was conducted to determine the isolated effect of the MERBS on ME retention in the RNZN.

The monetary benefits of any retention program are the avoided expenses for replacement and training of new personnel and the separation costs incurred for personnel leaving the military. There are also some less tangible benefits including higher experience levels, improved morale and increased flexibility. However, in order to remain objective, the focus of the study was placed on the monetary benefits achieved by the MERBS. There were two approaches taken to isolate the monetary benefits of the MERBS on retention.

The first approach was more subjective in nature and involved the use of a questionnaire taken by both the participants in the MERBS and their managers. The questions were tailored to evaluate the effectiveness of the MERBS and its isolated impact on the ME participants to stay in the RNZN. It was ultimately determined, based on these questionnaires, that forty one percent of the participants' decisions to stay in the RNZN were influenced by the MERBS. These forty one percent were then asked to rate the accuracy of their answer regarding the influence of the bonus payments on their decision to stay. The reply indicated they were ninety-three percent confident in the accuracy of their answer regarding the MERBS influence on their decision to stay (Oldham, et al. , 2002). The actual monetary benefit was determined by multiplying the participants' impact estimation of the retention payments by the estimated savings of retention for each of the 170 personnel participating in the MERBS. The estimated savings per participant and the detailed calculation of the ROI from this approach is shown in table 2.

The second approach involved a more objective approach using retention trend data. Predicted turnover

Table 2a: Determining ROI with the Participant Impact Estimation

Number of participants in the MERBS at the end of three year period	A	170
Estimated separation cost per ME leaving the service	B	\$4,260
Average replacement cost per ME leaving the service	C	\$105,133
Percentage of decision to stay influenced by retention payments	D	41%

Table 2b: Determining ROI with the Participant Impact Estimation

Number of participants in the MERBS at the end of three year period	A	170
Estimated separation cost per ME leaving the service	B	\$4,260
Average replacement cost per ME leaving the service	C	\$105,133
Percentage of decision to stay influenced by retention payments	D	41%
Percentage confidence in decision to stay influenced by retention payments	E	93%
Participants' estimation of retention payment's impact	$F = D \times E$	38%
Monetary benefits	$G = (B + C) \times A \times F$	\$7,066,789
Program costs	H	\$4,926,504
ROI from participant impact estimation	$\frac{G - H}{H}$	43%

Source: Author, 2009.

Table 3: Determining ROI with the Forecasting Method

Number of personnel retained attributed to MERBS initiative	A	73
Estimated separation cost per ME leaving the service	B	\$4,260
Average replacement cost per ME leaving the service	C	\$105,133
Monetary benefits	$D = (B + C) \times A$	\$7,985,689
Program costs	E	\$4,926,504
ROI from participant impact estimation	$\frac{D - E}{E}$	62%

Source: Author, 2009.

of ME personnel without the retention payments was estimated based on historical trends in both ME and non-ME personnel prior to the MERBS period. Based on previous trends of non-ME and ME personnel before the MERBS initiative, ME turnover averaged 5. 5 percent higher than non-ME personnel (Oldham, et al. , 2002). These data were used to determine an expected ME turnover without the retention payments. This was compared to the actual turnover during the MERBS period to determine an actual number of ME participants that were retained as a result of the bonus payments. Based on this comparison, it was concluded that seventy-three additional personnel were retained during the MERBS period than was predicted based on trend data without the MERBS initiative. The detailed determination of the ROI based on this second approach is shown in table 3.

It is interesting to note the similarities and differences between the two approaches. Both approaches determine program costs the same way. Even though the methods of determining monetary benefits are very different, the results are surprisingly similar. The participant impact estimation yields a monetary benefit of approximately seven million dollars while the forecasting method yields a monetary benefit of approximately eight million dollars. However, these numbers do yield significantly different ROI for the MERBS initiative.

The question then becomes which approach is more valid? Both approaches are logical and defendable. The forecasting approach is more objective since it is based entirely on data and trend analysis. However, the shortcoming is that it does not entirely isolate the effect of the MERBS initiative from the other retention initiatives occurring simultaneously. It is plausible that most of the

to Contents, page 1

seventy-three additional personnel retained were a result of the MERBS, but it is certainly possible that other proposed initiatives played a factor. The participant impact estimation approach is clearly more subjective since it is based on responses from a questionnaire. Conversely, it does better address the isolated effect of the MERBS initiative through the inclusion of specific questions in the questionnaire. Unfortunately, ROI methodologies based on nontraditional methods cannot always be purely objective. This is what makes it so challenging in the public sector. The objective is to develop a credible methodology, which is what was accomplished here. Therefore, both methodologies are valid as long as their shortcomings are kept in mind.

One lesson learned in the case of the RNZN relates to data collection. Ideally, the decision to perform an ROI determination is made prior to program implementation. This way data required can be determined and recorded while the program is taking place. For this case, the decision to determine the ROI for the MERBS was not made until after the program was implemented and well under way. Therefore, many data collection opportunities were missed. The recommendation of the study is to consider ROI evaluations and associated data collection needs during the development phase of program initiatives if possible (Oldham, et al. , 2002). Another lesson learned from the case was the extent taken to keep the ROI methodology simple. The RNZN could have attempted to determine monetary values for more complex but less tangible benefits. For example, they could have attempted to determine the monetary benefit of the increased operations tempo or increased experience levels available because of the higher retention rate. However, it is easy to imagine that this would involve potentially long and complex mathematical formulae and additional subjectivity. By avoiding these attempts, the ROI methodology is easier to understand and more credible. In this example there is a service center, a shadow price, and a quasi profit measure can be calculated.

United States Postal Service

The United States Postal Service used an Economic Value Added (EVA) program to determine their ROI from 1996-2002. EVA was calculated by determining the net operating income and subtracting a fee proportional to the cost of the assets used to produce that income (United States Postal Service, 2004). This difference represented a positive net cash flow that added financial value to the post office. A higher EVA indicated a more efficient use of assets. Consequently, senior post office executives were rewarded for performance at the USPS based on this figure. This provided financial incentive for post office employees to seek out new and better ways to improve efficiency within the organization. This program was credited with contributing to the $3. 5 billion in net income earned by the USPS from 1996-2000 (United States Postal Service, 2004). However, amid strong controversy relating to the calculation of EVA program incentive bonuses, the effort was abandoned in 2002.

The USPS 2003 Annual Report cites the continued use of ROI criteria for capital venture decision-making. Unlike many public sector organizations, the USPS is one of the few public sector companies that generate revenue. This facilitates using the traditional method of calculating ROI. The USPS continues to invest in automation equipment to reduce personnel work hours in mail processing and delivery (United States Postal Service, 2003). The cost savings realized from automation is then compared to the cost of acquiring the required equipment for making procurement decisions. The USPS also uses a Cash Flow/Capital Expenditure (CAPEX) ratio as a benchmark for assisting in making capital purchase decisions (United States Postal Service, 2004). The additional yearly cash flow to operations is compared to the yearly cash outlays to support the project. This helps to determine the attractiveness of the proposed project and the required need to borrow funds to support it.

The basis for determining economic value added at the USPS had significant flaws. The idea of subtracting additional costs from increased benefits to determine value added from a given investment is sound. However, the application of the EVA Variable Pay Program was inconsistent when overall USPS performance is considered. For example, the USPS lost $199 million in fiscal year 2000 but still paid out over $280 million in performance bonuses (Lexington Institute, 2001). As have observed during the U. S. federal government bank and insurance company bailout as part of the response to the financial stress conditions of 2008-2009, some critics deemed it improper and unethical for organizations that lose money to pay out significant sums of money in performance bonuses. In fact, the reason USPS lost money in the year in question was due to their bonus payouts. Of course, USPS is not supposed to earn a profit; their objective is to break even. Still, ROI measurements including those employing EVA appear not credible if they indicate positive results when other metrics such as negative net income indicate the contrary.

The USPS uses simple and intuitive methods for determining ROI for evaluating investment decisions. The use of cost savings as a basis for determining ROI is a commonly used method for public sector organizations. This is because many procurement decisions made by public sector organizations involve investments that will ultimately improve efficiency. If these efficiencies are able to be quantified, they can be used as a basis for comparison to the required capital expenditure to determine an ROI. The CAPEX ratio used by the USPS to evaluate capital purchase decisions is also an intuitive way to determine an ROI. Comparing cash flows to required capital 20 expenditures is very similar to a net present value calculation commonly used in the private sector to evaluate investment alternatives. It is important to emphasize that the USPS does generate annual revenues which lends itself to the use of traditional ROI metrics. This is uncommon among most other public sector organizations.

U. S. Navy Dental Corps

In response to the United States' Chief of Naval Operations' call for better decision-making tools, the Navy Dental Corps (NDC) has developed a simple metric to determine ROI at the branch clinic level. Captain York, the Navy representative at the Tri-Service Center for Oral Health Studies, spearheaded the effort to determine a practical method for defining NDC's return for investment dollars. While not using the traditional method of ROI that compares earnings to assets, this effort provides an easily understood metric to quantify performance at the branch clinic level.

NDC's ROI formula compares a branch clinic's quarterly output, defined as Dental Weighted Values (DWVs), to its required investment in funding (APF) and military labor (Milab). The formula is as follows (Mitton 2004):

$$\text{ROI}=\frac{(\text{DWV}\times 100)-[(\text{APF}+\text{Milab})\times 0.25]}{(\text{APF}+\text{Milab})\times 0.25}$$

Branch APF is the operation and maintenance funding allocated to the clinic. Both Branch APF and Annual Branch Milab are converted to quarterly values to determine quarterly ROI. DWVs and Annual Branch Milab are determined through separate data collecting programs described next.

DWVs are determined by input from the branch clinic into a program known as DENCAS. The clinic enters the different procedures performed on a given day using American Dental Association (ADA) procedural codes, known as Common Dental Terminology (CDT) codes. CDT codes are converted into DWVs that are essentially equal to one hundred dollars worth of dental services. This result is multiplied by one hundred to convert the DWVs directly into dollars for use in the ROI formula. Annual branch clinic labor is determined by the collection of data into the Medical Expense and Performance Reporting System (MEPRS). Branch clinic employees specifically document their hours worked performing a variety of individual tasks on MEPRS sheets. Different tasks such as various medical duties, training, and even leave/liberty times are documented. These data are correlated at the comptroller level to determine military labor hours and is converted into a dollar figure based on the rank and rate of the military employees working at the clinic.

The Navy Dental ROI formula discussed above is no longer used for three reasons. First, there is significant skepticism regarding the quality of the data being tracked for use in the calculation of ROI. Specifically mentioned was the inaccurate data collected by MEPRS. Many dental employees failed to log their hours on a daily basis; instead, they would record their hours on a weekly or monthly basis. This brings the accuracy of the type and number of hours into question due to the delay time in recording. Many times employees would wait until the end of the month and simply log eight hours of work arbitrarily for each day. Commander Mitton, from the Navy Bureau of Medicine and Surgery in Washington DC, referred to this popular method as "logging straight eights" (Mitton 2004). It is also difficult to use this formula for comparison. Different branch dental clinics may be responsible for different operating costs. For example, some of the clinics are responsible for paying their rent and utilities while other clinics are provided with these resources free of charge directly by the base command. This directly affects the amount of Branch APF the clinic would receive and, consequently, affected the results of the ROI formula.

Finally, the Navy Dental ROI formula does not include many of the cost elements required to staff and operate a branch dental facility. For example, large expenses such as the cost of training Navy dentists and dental technicians are not included. Other large costs such as accession bonuses for dentists and depreciation expenses for major equipment are also not included. Therefore, ROI for the Navy branch dental clinics needed to be more adequately defined. As a consequence, Captain York developed a more robust formula to be used in calculating ROI for Navy Dental clinics. Although similar to the previously discussed formula, it also includes many of the lacking cost elements. ROI is calculated as shown below (York 2004):

$$\text{ROI}=\frac{\text{Production Value}-\text{Cost of Production}}{\text{Cost of Production}}$$

Production value is determined similar to the DWVs calculated in the original formula. The cost of production, however, includes significantly more cost elements; these include system costs, which are the allocated training costs to the particular dental clinic from the dental training pipeline. These were not included in the original formula but are real costs burdened by the NDC and should be included. These system costs are divided between all the clinics proportionately based on the number of dental technicians and dental officers employed at the clinic.

There is still significant variation between the branch dental clinics use of the improved ROI formula. Table 4 compares the ROI calculated by the above formula for all the Navy branch clinics. Note that the ROI varies from approximately negative ten percent to over one hundred percent. These variations are not due solely to differences in performance levels. For example, NNDC is responsible for the costs of the Naval Dental Postgraduate School. Due to this fact, the NNDC ROI includes the impact of manpower costs and low productivity of the student-body significantly reducing their ROI. Therefore, even though this formula is more

to Contents, page 1

Table 4: ROI for Navy Branch Dental Clinics

Command	Cost	Production Value	Prod. Val. Cost	ROI
Great Lakes	$28,476,040	$60,746,220	$32,270,180	113.3%
Okinawa	$18,156,185	$30,022,556	$11,866,371	65.4%
Mid Atlantic	$29,003,564	$46,150,165	$17,146,601	59.1%
Parris Island	$10,81,353	$15,747,470	$4,937,117	45.7%
Camp Pendleton	$19,227,602	$26,027,670	$6,800,068	35.4%
Southwest	$34,282,076	$45,365,397	$11,083,321	32.3%
Southeast	$19,366,135	$24,946,372	$5,580,237	28.8%
Gulf Coast	$14,541,505	$18,741,273	$4,199,768	28.9%
Camp Lejeune	$17,474,322	$22,031,239	$4,556,917	26.1%
Europe	$12,454,115	$15,642,759	$3,188,644	25.6%
Pearl Harbor	$8,796,897	$9,973,433	$1,176,536	13.4%
Far East	$13,896,531	$15,376,018	$1,479,487	10.6%
Northeast	$11,073,148	$12,213,250	$1,140,102	10.3%
Northwest	$9,987,892	$9,711,583	-$276,309	-2.8%
NNDC	$32,151,242	$29,096,131	-$3,055,111	-9.5%
All NDCs	$279,697,607	$381,791,536	$102,093,929	36.5%

Source: Author, 2008

robust, one must consider more than just the final ROI output to fairly compare commands.

CONCLUSIONS

There are several challenges ahead for the development and implementation of notional return on investment. First, the subjective nature of the value-added factors will require senior level management buy-in to strengthen the formula's credibility. Phillips and Phillips point out three audiences which must buy in to the ROI process for it to be useful. They include the practitioners who are responsible for implementing the formula and held accountable by the results, senior level management who hold the practitioners accountable for these results, and researchers (Phillips and Phillips 2002). The NROI formula will not be used by the practitioners if it will not hold weight with their managers. This is true of any new process initiatives in the work place. Segregation of responsibilities with regard to data collection and evaluation must also be achieved to strengthen the credibility of the results. Also, the NROI process should be implemented at the inception stage of a project instead of during or after the project is in progress.

The NROI development methodology of defining how an organization adds value via its products or services can be applied to any public sector organization that is willing to go through the necessary steps to define how they create value in the products or services they provide. Several public sector organizations have already successfully implemented ROI metrics as one benchmark to demonstrate their performance. However, due to the distinct public sector outputs, this endeavor is often times challenging. Private sector organizations typically have the common goal of producing profit. Conversely, public sector organizations have a myriad of different goals most of which do not include the generation of profit.

This paper has probed the efforts of diverse institutions in the quest for a workable method of determining ROI in the public sector. Past efforts provided valuable lessons learned which were identified and discussed. The United States Postal Service was notably successful, but perhaps only because they function much more like a private company than is the norm for public sector entities. In Australia, the bold effort to develop a "value added" approach within the entire government budget process made some progress, but fell short primarily due to a focus on analysis lacking a quantifiable formula for determining the output. The Royal New Zealand Navy enjoyed relative success but was hampered in some ways by an inability to screen out other influences besides the bonus scheme for retention of marine engineers. The United States Navy Dental community effort was largely successful in the development of an ROI methodology, but fell short in their ability to use the output since they lacked trust in the method developed for determining some of the inputs.

Perhaps the most important lesson learned throughout this process was that the scope of such an effort must remain focused to be successful. Once NROI formulae are established at the project level and successfully demonstrated, an expansion of utilization may ultimately lead to the availability of ROI data for senior management decision making.

REFERENCES

Upper Saddle River, NJ: Prentice Hall, 2001)

Ray HG, Eric WN (2003). *Managerial Accounting* (San Francisco: McGraw-Hill Irwin, 2003), 542. It may be noted that private sector finance researchers sometimes refer to this method as "accounting ROI" (or Return on Assets -- ROA).

Richard B, Alan JM, Stewart CM (2004). Fundamentals of

Corporate Finance, 4th ed. (New York: McGraw-Hill/Irwin, 2004), 450-451

Phillips P, Phillips JJ (2002). Measuring ROI in the Public Sector, (Alexandria, VA: ASTD, 2002), 4.

Anthony E, Boardman et al (2001). Cost Benefit Analysis: Concepts and Practice, (Upper Saddle River, NJ: Prentice Hall, 2001).

Phillips P (2002). The Bottomline on ROI, Series Ed. Jack J. Phillips, (Atlanta: CEP Press, 2002).

Hawke A (2000). "What's the Matter – A Due Diligence Report, " (Address, Defence Watch Seminar, 17 Feb. 2000).<http://www. defencfor exampleov. au/media/2000/sec1. html [16 Nov 2004].

Australia Department of Finance and Administration (2003). "Value for Money Considerations, " Procurement Guidance, 24 June 2003 <http://www. financfor exampleov. au/ctc/toolkits/valueformoney/ value_for_money. html> [18 Feb. 2004].

Australia Department of Defence. Defence Procurement Review 2003. GPO, 2003.

Royal New Zealand Navy (2004). "What the Job Involves, " Marine Engineering Officer, <http://www. navy. mil. nz/future/role_detail. cfm?role_id=2> [16 Nov. 2004].

Oldham et al. (2002). "Calculating the Return on Investment of the Royal New Zealand Navy's Marine Engineering Retention Bonus Scheme." Measuring ROI in the Public Sector. Ed. Patricia Phillips. Series Ed. Jack J. Phillips. (Alexandria: ASTD, 2002) 132.

Oldham et al. (2002). "Calculating the Return on Investment of the Royal New Zealand Navy's Marine Engineering Retention Bonus Scheme." Measuring ROI in the Public Sector. Ed. Patricia Phillips. Series Ed. Jack J. Phillips. (Alexandria: ASTD, 2002) 138

Oldham et al. (2002). "Calculating the Return on Investment of the Royal New Zealand Navy's Marine Engineering Retention Bonus Scheme." Measuring ROI in the Public Sector. Ed. Patricia Phillips. Series Ed. Jack J. Phillips. (Alexandria: ASTD, 2002) 145

Oldham et al. (2002). "Calculating the Return on Investment of the Royal New Zealand Navy's Marine Engineering Retention Bonus Scheme." Measuring ROI in the Public Sector. Ed. Patricia Phillips. Series Ed. Jack J. Phillips. (Alexandria: ASTD, 2002) 141

Oldham et al. (2002). "Calculating the Return on Investment of the Royal New Zealand Navy's Marine Engineering Retention Bonus Scheme." Measuring ROI in the Public Sector. Ed. Patricia Phillips. Series Ed. Jack J. Phillips. (Alexandria: ASTD, 2002) 144-145

Oldham et al. (2002). "Calculating the Return on Investment of the Royal New Zealand Navy's Marine Engineering Retention Bonus Scheme." Measuring ROI in the Public Sector. Ed. Patricia Phillips. Series Ed. Jack J. Phillips. (Alexandria: ASTD, 2002) 146

United States Postal Service (2004). "Income-Economic Value Added, " United States Postal Service 1997 Annual Report, 7 Nov. 1997 <http://www. usps. com/history/anrpt97/management_analysis8. htm> [02 Mar. 2004].

United States Postal Service (2003). "Setting the Record Straight, " USPS Public Affairs and Communications, 31 July 2002 <http://www. usps. com/news/fyi/lex_response. htm> [13 Mar. 2004].

United States Postal Service. 2003 Annual Report. GPO, 2003. 25

USPS, 2003 Annual Report, 24

Lexington Institute (2004). "U. S. Postal Service Set Productivity Bar Low to Trigger Management Bonus Payouts This Year……even as the Service will lose more than $1 billion, " Postal Reform, 17 July 2001 <http://www. lexingtoninstitute. org/postal/productivity. asp> [09 Nov. 2004].

R. H. Mitton, Navy Bureau of Medicine and Surgery. Telephone Interview. 17 Mar. 2004.

York C (2004). "FY 00 Cost Value Model." Excel Spreadsheet. 17 Mar. 2004.

Phillips P, Phillips JJ (2002). Measuring ROI in the Public Sector, (Alexandria, VA: ASTD, 2002), 12

Prime Journals of Business Administration and Management (BAM)
ISSN: 2251-1261
Vol. 2(1), pp. 412-421, January 4th, 2012
www.primejournal.org/BAM
to Contents, page 1

Review

The Grameen Bank as an NGO – Successful or not?

Kazi Abdur Rouf

Social Economy Centre, University of Toronto.
Ryerson Entrepreneur Institute
Ted Rogers School of Management, Ryerson University, Toronto.

Accepted 21st December, 2011

The paper discusses the origin of GB and other NGOs in Bangladesh, and their impact on social development. Secondly, it also examines the extent to which credit helps women to be empowered; free from their male partners' dominance/violence; and are they able protest/fight against fundamentalism in Bangladesh. Thirdly, it discusses GB's management system, participatory rural approach (PRA), democracy, and accountability to citizenship. Fourthly, the essay looks at the commercialization of Grameen- phase-2 and its effects on advocacy for citizenship development. Lastly, it discusses the replication of the Grameen Bank model, and its future dimensions. From the research it appears that the Grameen Bank and other NGOs are playing a very important role in socio-economic development in Bangladesh although many of them have credit programs. Even though there are some criticisms about the controversial roles of NGOs, the widespread existence of GB and other NGOs involved in social development programs cannot be ignored. They are well able to improve many people's living standards. Although the Grameen Bank is firm to its credit delivery program it could initiate a pilot project for lobbying for women's equality and rights. This initiative can be achieved through involving people in civic engagement programs.

Key words: Grameen Bank, Microfinance Institutions (MFIs), NGOs, poverty, sixteen decisions, social development and women empowerment.

INTRODUCTION

Poverty, unemployment, hunger, malnutrition, illiteracy, gender discrimination and diseases are the major concerns in Bangladesh. It is a densely populated country, and therefore, it is hard for the government to provide adequate food, clothes, shelter, health, and education for its citizens. This failing has resulted in the emergence of the Grameen Bank (GB) as an NGO in Bangladesh. The Bank's high performance in poverty eradication has attracted donor agencies and caused them to shift their programs from the government to the NGOs because of the failure of the former to provide sufficient services and input to the countries destitute situation. Development through Grameen Bank and other NGOs is the new orthodoxy in Bangladesh. However, GB's main programs involve economic determination and it has recently become more profit driven. Hence, the question is how much are they actually contributing to the development of the country? Therefore, this paper is examining and analyzing the Grameen Bank to discover what it does well and where it needs to improve. The paper will examine which lessons can be taken from the Bank and applied elsewhere. This paper will also relate and examine the plan and actions taken by NGOs for social development in Bangladesh. Finally, it will propose solutions to the problems that the Grameen Bank faces.

First the paper will discuss the origin of GB and other NGOs in Bangladesh, and their impact on social development. Secondly, it will examine the extent to which credit helps women to be empowered; free from their male partners' dominance/violence; and are they able protest/fight against fundamentalism in Bangladesh. Thirdly, it will discuss GB's management system, participatory rural approach (PRA), democracy, and accountability to citizenship. Fourthly, the essay will look at the commercialization of Grameen- phase-2 and its effects on advocacy for citizenship development. Lastly, it will discuss the replication of the Grameen Bank model, and its future dimensions.

According to Erwin A. J. Dressen (2000) non-profit/non-government organizations and volunteer sectors are those who are registered charities, but the profits are returned to the organization. Michael Hall and Keith G.

Banting (2000) mention that different scholars use various terminologies for the nonprofit sector. For examples, *Nonprofit*' is the language of economists that view nonprofit organizations as a residual category; sociologists call it the *Voluntary Sector*; and political scientists, the *Third Sector.* According to Aditee Nag Chowdhury (1989) NGOs have been described as associations formed on personal initiative by a few committed people dedicated to implementing development projects at the grassroots level. They work outside the government structures but function within the legal framework of the country. GB was registered under special ordinance in Bangladesh in 1983 as a NGO, but has to follow central bank policies.

In the 1970s Bangladesh faced severe famine, hunger, starvation and poverty due to various economic, social and political crises. The rural poor in need of credit could not receive loans from the banks. They were outside the orbit of the banking system. In 1979, Professor Muhammed Yunus initiated the Grameen micro credit pilot project designed to address the economic crisis aiming to (1) extend banking facilities to poor people, especially women, and create self-employment opportunities, (2) eliminate the exploitation of the moneylenders, and (3) empower women through engaging them in income-generating activities. In October 1983 the Grameen Bank project was transformed into an independent bank by a government ordinance (P.O No. 26, 1972). There were several other NGOs which were officially recognized in the 1970s as well.

The first NGO activities started through Christian missionary groups. By 1793, the famous British missionary, William Carey, had arrived in Calcutta, India and began social welfare activities through the missionary-run school setup all over Bengal (Nurruzaman (1994), Abdur Karinm Khan (1981)). In the post 1971 period, however, there has been a mushrooming growth of NGOs in Bangladesh. Initially they rebuilt, provided relief, and rehabilitated the war-ravaged country. In the late 70s many NGOs were involved in eradicating poverty, malnutrition, illiteracy and setting up health programs. There has also been a tremendous increase in the number of NGOs in Bangladesh. Presently, there are 22,000 local and international NGOs registered in Bangladesh (Rahman, 2006: p. 454). According to Chowdhury Omar Faruque they provide services to nearly one-fifth of a population of 130 million.

There are some NGOs which have been successful in the development of Bangladesh since the 1980s such as the Bangladesh Rural Advancement Committee (BRAC), Proshika Manobik Unnayan Kendra (PROSHIKA), Association of Social Advancement (ASA), and CARE. They are best known for their multi-purpose development activities related to socioeconomic development in rural Bangladesh. For example, BRAC covers a whole range of activities for the rural poor, such as education, income-generation through poultry and livestock rearing, fisheries, forestry & handicrafts, just to name a few (Haque, M Shamsul 2002). However, it has been facing a whole lot of coercion and conditionality from the national government. Additionally these NGOs, including BRAC are severely dependent upon donor funds. In contrast, Grameen Bank gained popularity partly because it has become a self-sustainable organization as it works to alleviate poverty. The bank began its nationwide micro credit program to provide collateral free small loans to the landless poor. The present issue is how much it now contributes to social development which is not only concerned with economic enhancement.

According to the UN, social development means the integration of economic growth with social welfare policies leading to a higher standard of living for all. According to David Hollister (1982) social development is "the process of planned institutional change to bring about a better fit between human needs, social policies and programs. Based on the above definition, social development has several components including the eradication of poverty, expansion of education, and health for all and the assurance of fairly equitable distribution of income. My question is to what extent the Grameen Bank has, as a micro financing Institution (MFI) in Bangladesh affected changes in the level of social development?

GB has an intensive socio-economic program called the Sixteen Decisions, established in 1983, that helps to promote a peer support lending system among women. This creates social collateral which can be used by the Bank in lieu of material collateral. The Sixteen Decisions provide women with an opportunity to meet with each other and discuss various social development issues like health, education, environment, agriculture, and village events as well as pay installments. The Sixteen Decisions represent the social development agenda. Grameen centres are places for creating social solidarity among women. Professor Yunus says that, “Each center tries to ensure that all its members are guided by the Sixteen Decisions in their daily life” (Yunus, 1997, p.19). However, Rahman (1999) interprets this collective social collateral which should enhance social solidarity, as a strategy for ensuring high repayment rates, and which actually escalates violence against female borrowers (p. 72). While it is true that the Grameen Bank targets women more than men; it is with this strategy that GB directly channels loans to the poorest thus helping to improve their living standards. Along with providing credit, GB offers guidelines to members for codes of conduct and activities aimed at improving their social and financial conditions. It provides information to women concerning maternal health, nutrition, and childcare to generate a demand for basic health care services.

There is pervading evidence to be found in different credit review and NGO literature that GB has an overall positive impact on the lives of the poor in terms of

to Contents, page 1

creating higher income and self-employment opportunities as well as poverty alleviation. For example, the World Bank study (1998) indicates that 5 percent of Grameen Bank households rose above poverty each year. This GB micro credit system has had a great impact on extreme poverty (Khondaker 2003). It can play a vital role in attaining Millennium Development Goals (MDGs). In the words of Kofi Anan, UN Secretary General, “microfinance has proved its value, in many countries, as a weapon against poverty and hunger. It is recognized as a development model for income generation, self-employment and empowers disadvantaged women” (Grameen Dialouge-60, 2005). However, there is still an overwhelming amount of negativity surrounding the effects of NGOs and organizations like the Grameen Bank.

GB’s special feature is that 94% of its borrowers are women and they are in first place. The Bank encourages borrowers’ participation in the decision making process. It has easy policy guidelines for the staff. It uses local language and disburses loans in a very simply way. It has gained some concrete experience in its 28 years of operation such as: micro-credit being a very effective instrument to increase poor people’s income. Professor Muhammed Yunus, (2002) argues that welfare or handouts cannot help poor people. The Grameen Bank’s credit creates opportunities for poor women in Bangladesh to increase their income and overcome poverty.

Despite the activities of GB in Bangladesh with it main agenda of poverty alleviation, poverty still persists as a major problem. According to the Asian Development Bank (ADB) report, Bangladesh remains one of the poorest countries in the world with a per capita income of approximately US$ 337 in 1998 and with nearly one half of the population living below the poverty line (*The New Nation*, Dhaka, 02 December 1999). The pace of poverty reduction has been very slow. Another study shows that "at least 67.5 million rural people live in absolute poverty and of these 30 to 46 million exist in extreme, hard core poverty” (Cathrine H. Lovell, 1992: p. 11). There is no doubt that there has been a significant decline in poverty which has come down to below 50%.

My research uncovered two opposing views on NGOs. One is the pro-NGOs and the other, the anti-NGOs. The pro-NGOs view holds that NGOs are more effective than the government agencies because they operate at the grass roots level and work closely with the target groups that make them known to each other. Hence, they can easily identify the root of the problems, which could help them to take accurate action. They also argue that NGOs not only work for social development, but help people by providing access to modern technology, and education that contributes to social development (World Bank, Dhaka: 1994). However, there is also evidence to show that MFIs are exploiting the poor while doing business with them.

Although NGOs are non-profit organizations, the majority are exploiting the rural people (and especially women) through micro-credit programs. *The Daily Independent* newspaper in Bangladesh published an article that revealed an analysis of micro-finance data collected from 495 NGOs in the country. The results showed that about 95% of NGOs are directly involved in such suppression. It was also noticed that 95 per cent of NGOs charge interest rates within the range of 12 and 25 per cent. I contrast, even the largest NGOs – ASA, Proshika and Grameen Bank – only charge 15 per cent, 12.5 per cent, 18 per cent and 20 per cent respectively *(The Independent,* Dhaka: *December 21, 1999).*

With the anti-NGOs view, scholars assert that although NGOs are non-political and non-profitable organizations, their main aim seems to be political and profit driven rather than towards social development. According to Aditee Nag Chowdhury (1989) NGOs have cultivated a self-image that they are "development partners" working for removing poverty. Most NGOs do not properly use their funds. In fact, 70% of the NGOs funds and resources are spent on the salary and allowances of foreign employees, experts and consultants. According to Dr. Khaliquzzaman, "these NGOs fashionably named projects: poverty elimination projects, target group project, for example, have not brought about any changes in the economic structure in Bangladesh” (Nuruzzaman 1994: p. 8.). Thus the conclusion drawn was that NGOs have done very little for eradicating poverty from the country. In such situations GB is exceptional because its various studies provide affirmative information about poverty eradication; but there is still conflicting information found in the research regarding women’s empowerment.

Women’s empowerment is the process of gaining power over their lives, in terms of increasing welfare and reducing subordination to men through the expansion of their choice and voice in the decision making process. According to Lamiya Karim (2001) the Grameen Bank has provided the world’s financial community with the seductive information that: the poor are credit worthy and the Bank goes to the doors of the poor. But the high repayment rate 99% does not tell the reader how money is recovered from its poor borrowers. She finds that money is often recovered through intimidation, force and violence against poor female members (p.95). Simeen Mahamud (2004) asserts that although women have good access to credit, the norms and practices put them at a definite subordinate position to men, which means they have little power in the decisions that shape and render them vulnerable in the society. In this situation advocacy/lobbying programs can play an important part in fighting against patriarchy in the society.

Now I will discuss how GB’s credit provisions perpetuate violence against women. GB believes that income-generating activities by women could create the potential to enhance their power and status at home.

However, several studies indicate that although GB has contributed to poverty eradication and created self-employment, its activities have not yet been able to empower women to fight against patriarchy (Rahman, (1999), Goetz and Sengupta (1996), Karim (2001), Isseriles (2003), and Mahamud (2004).

Instead, loan liabilities raises tension and anxiety within borrowers' households and increased violence against women. For example, Rahman's (1999) study reports that 70% of borrowers confessed to increased violence and aggressive behavior in their household because of their involvement with the bank (p.74). "Men are users of more than 60% of women's loans (p.75). Karim (2001) finds that ninety percent of the loaned money went to men. She remarks that women are the bearers of the credit although not its users. The NGOs use women's social vulnerability and powerlessness but do not support them in the fight against patriarchy (Karim, p.100).

The same conclusion was reached by Isserles (2003) where he mentions that although GB has seen many success stories of women who have been emancipated through micro credit, he cautions that many women are still dominated by their husbands and do not actually reap the benefits of the financial gains of their investments. In many cases their husbands receive the credit (41). Muhammed Yunus admitted to some defects of the bank's lending practices in addition to the rigidity of the system. Therefore he made several fundamental changes in its operations in 2000 and called this new stage Grameen Bank Phase -11. He stated that, "Poor people are not trouble makers for the institutions, rather, the design of the institutions and rules make trouble for them' (2002). Now the questions raised to the GB are:

(1) do micro credit projects help to empower women or increase their vulnerability,

(2) does this credit lead to increased tension and violence for these same women?

In reference to Evans and Shields the Third Sector does good works. He suggests third sector organizations should lobby a 'mediation' role in society that can help people to build citizenship. According to them the third sector organizations are a central part of civil society. Civil society relates to the sector of 'space between public life and private life'. However, non-profit organizations are controlled by the state through service contracts. Recently, in Bangladesh, a few small third sectors/NGOs like *Proshika, Nijara Kori*, and *Unanyan Nari Bikalq* began to refocus their activities towards supporting civil society; however, the government placed injunctions on some of their activities causing them to malfunction. In addition, Islamic fundamentalists are strongly against of NGOs.

The Islamic fundamentalists (*Ulemas*) have raised strong protest against NGO activities and thus created tension in Bangladesh. They talk against Grameen Bank's credit program and others too. For example, *Ulemas* claimed that GB conducts interest earning investment activities which are prohibited (*Haram*) in Islam. They protest that the government must act against this revived form of imperialism and do so before the country loses its sovereignty and the nation, its Islamic identity (cited in Nuruzzaman, 1994: p. 9). The *Ulemas* also charged that great majorities of NGOs are allegedly engaged in missionary activities. Therefore their activities are seen as a threat, creating division in families which are the basic unit of Bangladeshi society (Islam, S. 1998: p. 84).

Hence there is persistent tension between NGOs and religious groups in Bangladesh. BRAC complained that 1,400 of its 20,000 schools are vandalized with a good number of them burned. The Grameen Bank also has complained about the 'fundamentalist's' negative attitude towards it *(Holiday-Dhaka 1994).* In the socio-economic field, despite the important roles played by NGOs, especially poverty eradication, some scholars argue that most NGOs are taking high interest from the poor. Therefore, according to them, NGOs are merchants, and tools of exploitation.

GB has autonomous board of directors who make its own policy and decisions, but its staff is paid. It is not a charitable organization. Its profits go back to the organization and to its beneficiaries. It is also not a volunteer organization. Studies have demonstrated that several NGO programs have had a significant impact on social & economic development in many marginalized households in the country. Due to this success, Bangladesh is home to some of the largest indigenous NGOs in the world. GB is the giant credit giving NGO in Bangladesh. According to Shamsul Haque (2002) Grameen Bank has been one of the most globally influential agencies regarding micro-credit in particular.

The Board of Directors is the highest policy making body, comprised of 13 members of whom nine are elected from among the borrowers. GB Board members are selected every three years. Muhamud Yunus is also a board member and the chief executive of the Bank. Various activities of the Bank are organized and implemented by four tiers of administrative set-up: branch offices, area offices, zonal offices and head office. All these layers are interconnected and follow a bottom up policy and decision making approach. The branch office is the lowest unit of operations of GB and it is located in a village. The branch organizes clienteles and disburses loans directly to borrowers. The area office supervises about 10 to 12 branches. The zonal office is located in the district headquarters. The head office is in Dhaka. The GB head office coordinates all field activities, and provides general guidance to field offices.

According to Kevin P. Kearns (1996) the accountability environment is a collection of forces – legal, political, socio-cultural and economic – that place pressure on organizations and people who work in them to engage in certain activities and refrain from engaging in others. They mandate certain organizational actions and prohibit

to Contents, page 1

others with a vast array of rules, procedures, reporting requirements, and sanctions by outside entities (p.29). But accountability is the bridge of answerability, responsibility and responsiveness (p. 38).

The Grameen Bank's plans, programs and strategies are not developed from the head office, but rather, field workers make customized decisions in the branch office that are appropriate and beneficial to the borrowers. Ninety three percent of the shares in the bank are owned by the borrowers and 7% remain with the government. All borrowers buy share certificates and thus become part owners of the bank. Staff at all levels is accountable and answerable to borrowers because in each branch there is a branch representative elected from the borrowers. Bank action plans are prepared and implemented through mutual dialogue among field staff and borrowers.

The Chairman of the board appointed from the government is usually a high government official. The Board of Directors sits every quarter of the year in which time they review bank activities, accounts and performance. They approve the annual budget and the annual plan of the bank. Here is the bottom up democratic decisions making process embedded in the board. Within the GB all activities are accountable to the board. Representatives from the branch, area, and zonal offices give their feedback to the board secretariat. The board reviews the field inputs in its meetings and makes decisions.

Although each NGO has its own agenda, strategies and programs for social development, most follow almost similar strategies, plans and actions in Bangladesh. For example, NGOs first motivate the isolated poor people to form groups to discuss their problems. After a period of conscientization, the needs for giving small loans for self-help projects are discussed by micro financing organizations. To do their jobs usually they follow some strategies for promoting the participation of beneficiaries: (1) *Localized Participation* attempts to increase the local people's abilities to participate in the social change process. The second is *Political Empowerment* which aims to promote involvement in broader social movements and increased political participation of those who are excluded from the national-decision making process. GB has been providing credit to poor people through groups which simultaneously work to enhance social development. However it has proved that it is difficult to attain a trickledown effect in the society where economic growth is concerned. Nevertheless, the Bank has some special strategies that help disadvantaged people to be included in its credit delivery system

Grameen bank is different from other commercial banks. Its actions are almost in direct opposition to other private banks. For examples, private banks give loans to rich people, GB gives loans to poor people; private banks only provide loans to men, while GB provides loans to women; schedule banks make large loans, while Grameen makes small loans. Commercial banks require collateral, but GB loans are collateral free. General bank clients go to the office, but GB workers go to the borrowers' doorsteps. GB provides loan through group formation whereas the transactions of general banks require no group formation. GB respects the local culture, norms, values, and religions. It works within the community as an insider. All the above are the strategies employed by Grameen Bank which has made it so popular to rural poor women in Bangladesh.

Grameen Bank targets and mobilizes the poor and creates social and financial conditions that help provide economic security through self-employment or otherwise, through the use of credit. The group lending, peer pressure mechanism helps to monitor and enforce contracts and aids in the screening of good borrowers from bad ones. GB mobilizes savings to develop good savings behavior among the borrowers. However, these savings mobilization schemes provide protection of loans against default, and provide an internal source of finance through stored funds. So savings are one kind of liquidity collateral. Moreover, field staff may deduct savings for delinquent loan adjustments.

The GB success in high volume lending, loan repayment rate and high profit is well-liked among development practitioners. Grameen Bank operates nationwide through 2,185 branches. The repayment rate has been highly satisfactory (99%) since 1979. The majority of patrons (97.9%) are female borrowers. The Bank serves a total of 6.7 million borrowers through 130, 000 rural landless associations in 70,370 villages in Bangladesh. The total loan disbursement has been $5.65 billion since its inception. Of them $5.00 billion has been repaid. Current borrower savings are $2.2 billion (Grameen Updates 2006). It runs on its own internal funds and investment incomes and that makes it economically sustainable. However, Isserles (2002) challenges that the Bank's high repayment rates do not indicate high efficiency; but he fails to point out the increase in the quality of women's lives.

Building capacity is about improving the quality of life for all. According to Sherri Torjman, building community capacity is a holistic approach to solving problems and improving the quality of life. It is built upon the intrinsic links among economic, social and environmental well-being. It is 'pro-people, pro-nature and pro-jobs' (Torjman, 1998: p.3). GB is working for poor people and its programs and policies are developed by the representatives of borrowers. However, it has no integrated development programs like literacy, family planning, rural infrastructural development, development of voting behaviors, citizenry skills development and efforts to work against patriarchy, for example. It only works for poverty through credit delivery to poor people. Thus GB's credit activities cannot significantly affect sustainable development unless it creates and implements holistic development programs. Sherri Torjman mentions seven strategies for sustainable development. These strategies are: (1) poverty reduction, (2) broadened concept of investment, (3) civic engagement, (4) problem-solving, (5) partnership, (6) leadership development and (7) celebration. Grameen Bank only has a poverty reduction strategy; unfortunately all other strategies are important for sustainable development and to build skills among the poor. These are also important for accountability and for transparency within the organization.

Kevin P. Kearns (2006) in his article “From Adequate to Outstanding Performance” mentions six propositions that can make an outstanding organization. Such outstanding organizations continuously adapt and refine their missions, vision and aspirations with change, time and situations and refine innovative and effective approaches to accomplishing their mission Outstanding performance begins with outstanding leadership. In the case of the Grameen Bank, such propositions have praxis at the ground level. For example, GB has changed its different plans, strategies, and revised its various loan products according to the needs of its borrowers. When the GB was established in 1979, there was only one loan product, but now it has housing loans, disaster loans, and education loans in addition to its basic loan. It has moved to profit making for its financial sustainability. This change, effected for organizational survival, Kearns calls mission drift (p. 3).

According to Jim Collins (2001) an organization could be excellent if it has a clear focus on goals, expectations, and accountabilities and does not achieve greatness overnight. In this perspective Grameen Bank can be in the category of an excellent organization because it has a clear vision of eradicating poverty among poor women. The leadership of Muhamed Yunus has been the guiding force for Grameen’s initiation and expansion. However, over time, Grameen Bank has institutionalized a decentralized management structure with a cadre of dedicated professionals that is operating without much of his involvement. Field staff is implementing the credit program. There are several departments in the head office including administration, accounts, audit, and MIS. These departments support and coordinate field activities. According to Collins, effective leadership connects with mission growth because the mission growth is a process of adapting and refining the organization’s mission, vision, and aspirations for expanded community impact. In Grameen Bank there is a strong leadership which drives the mission to growth.

To develop gender equality through women’s leadership development each year the center chief for GB community groups is changed, which is an opportunity to develop each borrower’s leadership qualities. Professor Yunus says that, “The same person cannot be elected twice until other members of the group have had their chance” (Yunus, 1997, p.13). It is a system through which the center develops poor women’s leadership in rural Bangladesh and raises their self-confidence (Chandler, 1993). However, Aminur Rahman (1999) discovered that one or two influential members had real control over the decision-making process of the center (p.74). Center chiefs hold *de facto* power as they decide almost every issue of the center. Perpetuation of such power relations in the loan centers is contradictory to the Grameen Bank ideology.

Armstrong and Mollenhauer (2006) discuss the culture of accountability within organizations. In their words, “A culture of accountability describes the critical success factors and indicators that need to be present in a highly accountable organization (p. 1). The success factors are strategic leadership, performance culture, and rigorous decision; abilities to generate reliable information, create an environment of innovation, access and manage risk, and the possession of clear values and ethics. Grameen shows such success in its work in Bangladesh. For example, in 1988 and 1998 there were some major floods which destroyed farmers/borrowers standing crops, resources and houses. GB borrowers were seriously affected by these two floods. In the periods following the floods, the Bank provided extra loans to its borrowers to rehabilitate their businesses. This disaster loan helped borrowers to recover their businesses, renovate their houses and cultivate crops. However, this disaster loan came with a 20% interest similar to the basic loan. Many journalists criticized it and compared GB to the East India Company in their dealings with the poor.

GB staff members are committed to implementing its credit programs for the benefit of poverty eradication and the creation of self-employment opportunities. Grameen engages its stakeholders and creates meaningful dialogue with the more critical parties such as the NGO Bureau of Affairs and the Ministry of Finance, and invite them to participate actively in helping to achieve its mission. The bank also collaborates with Grameen Trust to replicate its peer support lending technology to other agencies with the help of UN agencies, CIDA and USAID.

According to Sonia Opina, William Diaz and James O’Sullivan (2002) negotiated accountability must encompass involving the range of stakeholders from the initial inception of the program, through to its implementation, evaluation and finally to the dissemination of results. Good communication mechanisms are important tools in negotiating the accountability environment, in order to learn about and respond to the needs of the community. In negotiation processes, stakeholders need to be flexible in their win-win situation. However, GB is always looking for a win-win situation and that sometimes creates problems with its shareholders. For example, in 1983 GB wanted 100% of share capital to be owned by its borrowers, but instead, the government registered it with 60% owned by the government. After a long process of negotiation GB achieved 93% of shares for its borrowers, while the government retained 7% of capital shares.

James M. Ferris emphasizes the autonomy of NGOs to

to Contents, page 1

implement their own decisions. Unfortunately, the government often regulates the activities of nonprofit organizations, using them as instruments for carrying out public policies; in particular, the expansion of welfare state expenditures at the national level (363). In practice, there is really no absolute autonomy in the nonprofit sector, but rather all organizations are held accountable to the NGO Bureau. Although the government should monitor NGOs funding and activities, it is not acceptable to regulate and place conditions on them that might hamper their abilities to manage their programs effectively.

There are many NGOs in Bangladesh involved in adult literacy, maternal and child health, primary health care, nutrition, education, agriculture, relief and rehabilitation programs. For instance, BRAC itself has established more than 5,000 primary schools and kindergartens where about half a million students are enrolled. It is very interesting that government schools have 15% attendance, whereas, NGOs have about 90% attendance (Richard Holloway: 1998, p. 22). Consequently, the literacy rate has increased. The latest statistics shows that the adult literacy rate has increased over the last few years, especially female, from 16% to 38.1%. However, very few NGOs are involved in vocational skills development for the poor. Grameen bank and other MFIs are not involved in human skills development. Moreover, large NGOs do not develop their role in the area of environmental sanitation, which includes provision of fresh water and disposal of waste (Aditee Nag Chowdhury 1989: p. 202). GB and other MFIs can step in for environmentalism.

Recently the Grameen Bank shifted to a profit motive instead of retaining its original objects. This profit-making focus diverted the focus of employees to credit disbursement and repayments rates; simultaneously turning them away from social and citizenship development issues. This is despite the fact that GB had objectives of creating organizational solidarity and networks among the poor and bringing groups together in co-operative activities such as the joint effort solution of problems. This system has led to more opportunities for women to show leadership. However Anne Maria and Sengupta (1996) challenge this and infer that GB focuses on dispersal of credit and repayment but is unable to create solidarity among women. Rahman (1999) remarks that Grameen Bank has a really good objective, but now there is a gap between its original goals and field realities.

Drake and Rhyne (2002) concluded that instead of helping to eradicate poverty MFIs are involved in doing high profit business. Although Professor Yunus of Grameen Bank cautiously makes an important point in that the development finance community must continue to innovate in the areas of targeting tools, products, service delivery and operations in order to reach the full range of low income families. However, a focus on financial sustainability encourages MFIs to place a greater emphasis on making profit than on the well-being of clients as this has a negative impact on borrowers' lives (Goetz and Sen Gupta (1996) Morduch (1999) Rahaman (1999), Woller 2002). For example, GB made a profit of $6.03 million dollars in 2005 (Gramen Bank annual report 2005, p. 54). Although Grameen Bank claims to be a social benefit institution, it now operates on a profit motive as a private bank.

In 2000, GB reformed some of its products and introduced a new system called the Grameen Generalized System (GSS). Before 2000, all Grameen operations were termed Grameen Classical System (Phase -1) and after 2000 the same operation was renamed, Phase-2. In Phase-2, GB is more flexible and friendly to borrowers in loan transactions. The new system's main features are prime loan products including basic loans, housing loans and the higher education loan. Repayments can be made according to borrowers' income and loan size as per repayment records. Grameen had tested this new mechanism in the field before establishing it in all branches. However, this new initiative has challenged the bank to sustain work with the poor in the lowest realms of poverty.

One must consider why citizenry is important and what role the GB plays in citizenship skills development for low income women in Bangladesh. If we consider empowerment as 'the process of challenging existing power relations, and of gaining greater control over the sources of power then it gives conflicting information. For example the Kabeer (1998) study, *The Small Enterprise Development Program (SEDP) of Bangladesh* draws the conclusion that micro-credit programs have a positive effect, but it is largely men who benefit, because micro-credit has done little to move toward a situation of gender equality.

However, in Canada NGOs are diversified and are vehicles for the engagement of citizens and environmentalism. For example, the Ontario Coalition for Advocacy against Poverty (OCAP), Canadian Council on Social Development (CCSD) and Canadian Ethnocultural Association (CEC) are involved in advocacy programs, as well as in protesting, advocating and lobbying for human rights and women's rights. Campaign 2000 for the National Child Benefit (NCB) influenced the government to change some policies on child poverty and childcare. Although the Campaign 2000 is unpleasant, risky and controversial, it is successful in increasing child benefits for Canadian children. Likewise, Grameen Bank and other NGOs can make alliances and lobby for participatory democracy and equity in all public policies. Canadian lobbyist organizations are strong and popular in Canada in contrast to the weakness of Bangladeshi NGOs in lobbying for women's rights.

Ilan Kapoor (2002) mentions that the PRA campaigns for local knowledge and puts forth a methodology aimed at enabling local people to take power over their own lives. Ultimately these strategies can help form an understanding of the power and politics of the society. Participatory rural appraisal (PRA) is based on equitable and respectful partnerships and collaboration between clients and institutions with due attention to local knowledge (Chamber (1997, Crush 1995, Escobar 1995, Friedmann 1992). According to Amartya Sen (1999) it is a process of developing individual capabilities through education and skills in order to power individuals to fight for a better life, democracy and economic liberalization. Grameen Bank and all other MFIs are working closely with marginalized people at the grassroots level. This level of organizational development is an integral part of Grameen's credit program. It follows the Participatory Approach in working with the poor building from a location-specific context. GB believes that if group solidarity is strong, the poor can together form a socially and politically powerful entity. However, in practice GB borrowers are far behind to challenge the power structure of the society.

Robert D. Putnam (2000) says that if participation in political deliberation declines – if fewer and fewer voices engage in democratic debate – then politics will become shriller and less balanced (p. 338). Disadvantaged people will be further exploited in the public sphere. GB follows the democratic process by electing group leaders; however, it does not have programs in which borrowers can be involved in the legislative electoral process that can create a real democratic voice in the public sphere especially in politics in order to establish their rights in the society.

Therefore Grameen Bank and other NGO programs on Civic Engagement are very essential in Bangladesh to mobilize citizenry because strong and caring communalities start with the citizen as the base of their development. Civic engagement or broad participation in decision making is a prerequisite for achieving sustainable development. Professor Yunus realizes this and took the initiative to be involved in politics in 2007. He however did not continue his commitment to be engaged in politics. Other NGO leaders did not respond to his new mission. Even Grameen Bank does not take up the challenge of activism for citizenry.

The GB and other MFI's have been working in close contact with the rural poor. It is estimated that NGOs employ 200,000 young men and women as fieldworkers. The educated NGO workers have daily contact with the villagers, visiting them in their homes. In the 1980s the staff of Grameen bank spent more time in the centers discussing various social issues with group members. However, now the Bank workers are involved with different types of loan investments like basic loans and housing loans and collect them from the borrowers (Bernasek, 1992; Chandler 1993; Goetz and Sen Gupta, 1996). These activities, according to Hashemi, " this going to the poor breaks down some of the threatening distance between the educated NGO workers and the poor, that is so much a part of rural social stratification (Hashemi 1997b:4). According to Karim, "these activities in reality, introduce a new power elite into the existing dynamic of rural social relations (2001. p. 96).

In West Bengal popular political mobilization has successfully led to pro-poor economic reforms. The rural poor people were able to use local government as an alternative institutional channel for promoting their interests (Sengupta and Gazdar, 1998). Union council is a local government institution that has been involved in local political power, but is less powerful in helping the local people. Professor Yunus recently tried to draw the government's attention to focus on this local institution, but GB is not directly involved in lobbying for citizenry.

According to the Grameen Bank Annual Report of 2005, 58% of the families of Grameen borrowers have moved above the poverty line. The remaining families are moving towards the poverty line. However if we compare Grameen borrowers' electoral success with their economic elevation, we can draw an assumption that political achievement is not as strong as its economic improvement. For example, in 1997, out of a total of 4298 unions of Bangladesh, 7 GB members and 39 members of the family were elected to the post of Chairman. In addition, 1,073 family members were elected as members of the union council (Grameen Dialouge 2001, p.16). However, this data is not as strong as its poverty success statistics.

Recently, BRAC initiated a TUP program that re-engages social mobilization and advocacy programs. However it is on a small scale. Massive expansion of such programs can help to create a voice for the poor. GB can initiate such an advocacy program that can help the poor gain access to political institutions; and thus challenge and reform public policy.

Since 1989, Grameen Trust (GT) has played a role in promoting the Grameen type micro credit programs in Bangladesh and in different countries. It initiated a number of international dialogue programs where many people came from different countries to participate and to learn and share their experiences with Grameen loan programs. By June 2006, GT partners had collectively reached more than 3.14 million poor families with micro credit in 37 countries through 143 projects in Asia, the Pacific, Europe, Africa and the Americas (Grameen Dialogue, 2006).

However, a number of problems arise during the initial replication and implementation times. Some replication projects are not successful. For example, The Good Faith Fund in Arkansas initiated a Grameen type loan project in 1990. Although this project is popular it has since merged with the local credit union. However, Grameen type credit programs in India, Pakistan, Philippines, Myanmar, China, Nepal, and Bhutan continue to be successful.

From the above discussion, it appears that the Grameen Bank and other NGOs are playing a very important role in socio-economic development in

to Contents, page 1

Bangladesh although many of them have credit programs. Although there are some criticisms about the controversial roles of NGOs, the widespread existence of GB and other NGOs involved in social development programs cannot be ignored. They are well able to improve many people's living standards. The unfortunate truth; however, is that very few of them are involved in democracy and advocacy programs for citizenry skills development of the poor. Therefore, it is very important that GB and other NGOs include an alternative model of democratic development in their programs for strengthening awareness among the disadvantaged people. At the same time, government should not isolate NGOs from their development programs but rather, work in cohorts with such organizations. NGOs should also understand the local environment and culture. Furthermore, although the Grameen Bank is firm to its credit delivery program it could initiate a pilot project for lobbying for women's equality and rights. This initiative can be achieved through involving people in civic engagement programs.

Acronyms: Association of Social Advancement (ASA), Bangladesh Rural Advancement Committee (BRAC), Grameen Bank (GB), non-governmental organizations (NGOs), micro-finance institutions (MFIs), Millennium Development Goals (MDGs), participatory rural appraisal (PRA), Proshika Manobik Unnayan Kendra (PROSHIKA).

Abbreviations

Participatory rural approach (PRA)
Micro financing Institution (MFI)
Bangladesh Rural Advancement Committee (BRAC),
Proshika Manobik Unnayan Kendra (PROSHIKA),
Association of Social Advancement (ASA),
Millennium Development Goals (MDGs).
Asian Development Bank (ADB)
Grameen Trust (GT

REFERENCE

Abdul Karim Khan (1981). Christian Missions in Bangladesh: A survey. London: Islamic Foundation, (p. 4).

Aditee Nag Chowdhury (1989). Let Grassroots Speak: People's Participation, Self-help Groups, and NGOs in Bangladesh (pp. 18-19). Dhaka: University Press Limited.

Armstrong, R. and Mollenhauer, l (2006). "A Culture of Accountability. Ontario Trillium Foundation.

Bannerji, H. (1993). Popular Images of South Asian Women. Returning the gaze: Essays on racism, feminism, and politics. Toronto, ON: sister Vision Press.

Barnes, Barry (1995). 'Social Movements,' in The Element of Social Theory.. Princeton, N.J: Princeton University Press. pp.151-171

BRAC Research and Education Department (2004). "The Elites, Their Perspectives of Poverty, and the Gram Shahayek Committee: Some Early Explorations", CFPR/TUP Research Preview, No 7.

Cathrine H. Lovell (1992). *Breaking the Cycle of Poverty* Connecticut. Kumanan Press Inc.

Cited in Nuruzzaman (1994). "NGOs-The Web of New Colonialism: Aid Merchant Buying up Sonar Bangla", Impact International (London), July 1994, p.8)

Chowdhury, Omar Faruque (2005). "The Diversity of Volunteerism: Bangladesh Perspective"
http://www.un.org/dpi/ngosection/chowdhur.htm

Cited in Nuruzzaman (1994). "NGOs-The Web of New Colonialism: Aid Merchant Buying up Sonar Bangla", Impact International (London), July 1994, p.8)

Drake, C. and Rhyne, E. Eds., (2002). The Commercialization of Micro-Finance: Balancing Business and Development. Blomfield, C.T. Kumarian Press. Chapter 1,2, and 6.

De Tocqueville, Alexis (2000). 'The Influence of Democracy on the Sentiment of Americas,' chapters 1-7, in *Democracy in America*, .P. Mayer, ed., New York, Perennial Classcs.

Dressen A. J. (2000). What do We Know About the Volunteer sectors? An Overview, ' Research Paper, Statistics Canada.

Evans, B. Mitchell and Shields, John. (2000). 'Neoliberal Restructuring and the Third Sector: Reshaping Governance, Civil society and local relations. Department of Politics and School of Public Administration, Ryerson Polytechnic University.

Farid Hossain, "Mullahs Escalate NGO-Bashing", Holiday (Dhaka), April 1, 1994).

Fuglesang, A. and Chandler, D. (1889). "Participation as process: What we learn from Bangladesh" Dhaka: Grameen Trust.

Goetz, A. M. and Gupta, R.. (1996). Who Talks the Credit? Gender, Power, and Control over Loan Use in Rural Credit Programs in Bangladesh. World Development 24(1).

Grameen Bank At a Glance (2006). Dhaka: Grameen Bank.

Gramen Dialouge (2006). Grameen Trust, Dhaka.

Grameen Dialogue- 60 (2005). Grameen Trust, Dhaka.

Grameen Trust Monthly Report (June 2006). Dhaka: Grameen Trust.

Haque, M Shamsul. "The Changing Balance of Power Between the Government and NGO in Bangladesh." International Political Science Review 23 Nov-2002: 413-415

Hollister, D. (1982). "The Knowledge and Skill Bases of Social development", in Daniel S. Sanders (ed.), *The Development Perspective in Social Work* (Honolulu: University of Hawaii.

IDB statistics and Quantitative Analysis Unit, Internet Report.

Isserles, Robin, G. 2003. Microcredit The Rhetoric of Empowerment, the Reality of "Development As Usual", *Women's Studies Quarterly*, 3 & 4.

James M. Ferris(1993). “The Double-Edged Sword of Social Service Contracting: Public Accountability vs. Nonprofit Autonomy”. Nonprofit Management & Leadership, Vol. 3, No. 4, p. 336-376.

Kobayashi, Audrey (2000). ‘Advocacy from the Margins: The Role of Minority Ethocultural associations in Affecting Public Policy in Canada,’ in the Nonprofit Sector in Canada: Roles and Reltionships, Banting, Keith G., Ed., 229-261. MacGil-Queen’s University Press.

Kevin P. Kearns (1996). “Preserving the public Trust-. Managing for Accountability”. Jossey-Bass publishers, pp. 29-44.

Kevin Kearns (2006). “From Adequate to outstanding performance”. The Forbes Funds, pp.1-18.

Latifee. H. T. (2002). Grameen Trust Experience (1991-2002). Dhaka: Grameen Trust Publication.

Mahamud, S. (2004). Micro-credit and Women’s Empowerment in Bangladesh in “Attacking Poverty with Micro-credit, “. Eds. Salehuddin Ahamed and M. A. Hakim. Dhaka: Dhaka University Press.

Michael Hall and Keith G. Banting (2000). The Nonprofit Sector in Canada: An Introduction, ‘in the Nonprofit Sector in Canada: Roles and Relationships, Banting keith G., 1-28. McGill-queen’s University Press.

Mizan, A. N. (1994). *In Quest of Empowerment: The Grameen an Impact on Women’s Power and Status.* Dhaka: United Press Ltd.

Rahman, A. (1999). Micro-credit Initiatives for equitable and Sustainable Development: Who pays? World Development Vol.27, No.1. Elsevier Science Ltd.

Robert, P. (2000). “‘Democracy,’ in Bowling Alone: The Collapse and Revival of American Community”. New York: Simon & Schuster. pp. 336-349.

Richard Holloway (1998). “Supporting Citizens’ Initiatives: Bangladesh’s NGOs and Society”. London: International Technology Publications.

Ruth Armstrong and Linda Molenhaurer (2006). “A Culture of Accountability”. Ontario Trillium Foundation. Toronto.

Qazi Kholiquzzaman Ahmad (1988). *Development communication and grassroots participation: Bangladesh Case* (Dhaka: Bangladesh Institute of Development Studies,

Shahid Khanker (1998). “Fighting Poverty with Micro credit- Experience in Bangladesh.” New York: Oxford University Press Inc.

Karim, L. (2001) Politics of the Poor? NGOs and Grass-roots Political Mobilization in Bangladesh. American Anthropological Association.

Khandker, S & Latif,M. (1996). *The role of Family Planning and Targeted Credit Programs in Demographic Change in Bangladesh i*n World Bank discussion paper no. 337. Washington: The World Bank.

Shamsul, H. (2002). “The Changing Balance of Power Between the Government and NGO in Bangladesh.” International Political Science Review 23 (2002): 411-435.

Sherri Torjman. (1998). “Strategies for Carrying Society”. Caledon Institute of Social Policy Presentation. ISBN#1-859194-23-3. pp. 1-20.

Sonia Ospina, William Diaz, Lames O’ Sullivan (2002). “Negotiating Accountability: Managerial Lessons from Identity-Based Non-profit organizations”. Nonprofit and Volunteer Sector Quarterly, Vol. 31, No. 1, p. 5-31. http://nvs.sagepub.com/cgireprint/31/1/5

Syed Serajul Islam (1998). "Impacts of Technology and NGOs on Social Development: the Case of Bangladesh", (P. 57) *Journal of South Asian and Middle Eastern Studies*, vol. 21, no. 2, winter

Thornton, P., J. Devine, P. Houtzager and D. Wright. (2000). Partners in Development. A review of Big NGOs in Bangladesh.” Report commissioned by DFID, Bangladesh.

Yunus, M. (2006). Grameen Bank at glance. Dhaka.

Yunus, M. (2005). Grameen Struggling Program. State of the Micro credit Summit Campaign Report 2005. Washington.

Yunus, M. (2003). What is Microcredit? Dhaka :Grameen Bank,

“Talking about Charities 2004”, Muttart Foundation. Canada.

The Grameen Bank Ordinance. (1983). Dhaka.

The Morning Sun (Dhaka), April 27, 1993. For instance, it is reported that an NGOs Nijera Kori, supported by Christian Aid, "War on Want" is politically active.

The Independent (Dhaka), Internet Edition, December 21, 1999

*The New Nation (*Dhaka), Internet edition, September 11, 1996

The UN Research Institute for Development, *The Commission for Social Development* (New York: 1963). See also ILO, "Perspectives: World Summit for Social development: Impetus to Concerted Action." *International Labor Reviews*, 134, 2, (1995), PP. 243-258.

World Bank Annual Report (1994). Published by Association of Development Agencies in Bangladesh. Dhaka.

Wuthnow, Robert (1991) ‘The Voluntary sector: Legacy of the Past, Hope for the Future?’ in *Between States and Markets: the volunteer Sector in Comparative Perspective*, Robert Wuthnow, ed., Princeton, Princeton University Press. pp. 3-28.

Prime Journal of Business Administration and Management (BAM)
ISSN: 2251-1261
Vol. 2(1), pp. 422-425, January 4th, 2012
www. primejournal. org/BAM
to Contents, page 1

Full Length Research

Evaluating the impact of training and development on organisational productivity: evidence from Nigeria

Obasan A Kehinde

Department of Business Administration, Olabisi Onabanjo University, Ago – Iwoye, Ogun State, Nigeria. E-mail: obasankehinde@yahoo.com, Tel- +234-805-614-1961

Accepted 20th December, 2011

Relying on controversies surrounding the usefulness of Training and Development at promoting state of the art skills to cope with challenges of managing modern days organisations for optimum productivity, this study, thus becomes imperatives at investigating the impact of Training and Development on Organisational Productivity. To achieve this, the study made use of survey research design through administration of well-structured questionnaire administered on selected staff of one of the selected local government employees in Nigeria. Consequently, the objectives as well as the hypothesis were achieved and tested respectively using correlation approach. The analysis then revealed that there is significant relationship between Training and Development and Organisational Productivity. On this note, it is safe to say that the evaluation is positive and thus recommended that organisations should continuously train and retrain their staffs in order to gain maximum contributions from them towards the realization of organisational objectives.

Keyword: Organisation, training, development, performance, productivity

INTRODUCTION

Regular training and development of its employees is a pre-requisite obligation of a future orientated and goal seeking employer. A practice which entails impacting skills and knowledge on the employees in order to cope with the various phases and challenges of the business environment which determines the success and quality of a firm in meeting the customers demand, competition and opportunities that lies beyond their scope.

According to Subramony (2009) an employee has a great role in influencing the work outcomes of an organization and that constant training of workers is a determining factor in the attainment of required skills for better job performance and manpower development in organisations. In the same vein, Dale (2002) observed that employers who have the intention to increase their productivity and output must enrich their workers in terms of training and development opportunities to fulfill the employers' ambitions. If these employees see that the organisation provides room for them to realize their potentials, personal growth and career, they will be opted to remain and discharge their best performance, otherwise, these employees will likely leave to fulfill their aspirations elsewhere.

From the foregoing, Valle et al (2000) considered training as a critical function of maintaining and developing working capabilities in employees. It involves the process of teaching new and present employees the skills they need to perform jobs which consists of those activities that are designed to improve individual's performance in currently held job or one related to it while Holton (1999) defines development as the process used to advance an employee to the desired level of performance.

The major purpose of training and development is to eliminate performance deficiencies whether current or anticipated through an improved performance, especially of organisations with stagnant or declining rates of productivity. Hence, it becomes the responsibility of each organisation to design ways of enhancing workers skills so as to meet up with the set target of the organisation within the realm of over-growing challenges of the global standards as the society itself is dynamics.

The importance of training and development on employees and organisation's effectiveness cannot be over-emphasized, Hassan (2007) opined that firms strongly desire to promote values such as trust creativity and quality in their employees and for that, proper training and development is necessary while Brewer et al (2008) observed that work force development and transfer of training are important concerns of any organization. Hence Human resource practices such as employee selection, appraisal, training and development, compensation are the key point in the performance of an organization since they ensure the uniformity of the set goals and the eventual outcome.

The study seeks to examine the impact of training and development on the productivity of the employees' vis-à-vis organizational goal attainment in Ikosi-Ejirin local council, one of the 37 local development council areas created in Lagos State, Nigeria by the Bola-Tinubu led administration.

Literature review

Overtime, there are controversies in .the academic field on the nature and role of training and manpower development in organizational goal attainment. Similarly, studies on manpower training and development tactics is not new as it can be traced to the medieval period when Greek philosophers such as Democritus, Plato, Aristotle among others postulated about the relationship between management and human behaviour since such field is concerned with organizational activity aimed at improving the performance of individuals and groups in organizational settings.

Many countries recognize the importance of training and development especially with the emergence of the knowledge based economy. Most governments have recognized that effective human resource training and development is necessary to raise the standard of living for their countries. Many states are pursuing industrial development to improve organizational performance, but industrial development will not be possible without an adequate pool of the right people to do the right jobs as such, many states incorporated national manpower training as an integral part of their economic plans to achieve economic growth while such formulation taking into account the training requirements in different sectors.

Zhu et al. (2009) highlight general trends of HRM changes in terms of people management systems and illustrate the underpinning factors, for example traditional values and culture, historical evolution, political and economic changes, and characteristics of society, industry and firm in each country that determine the formation and reformation of management thinking as well as HRM policies and practices. Hence it is strongly believed that of all the factors of production available to any organization, human capital remains the most important in the 21st century, even with advent of advanced technology like the use of computers and automated machines. This belief has remained unchanged. In the same vain, this is necessary in order to improve the performance of people at work which is important as establishing the organization as an entity. The human capital serves as a source through which organizations could generate a competitive edge that cannot be easily reproduced.

There is practically no successful business strategy that would not put into consideration the impute of human relation since there is no strategy that can be introduce at the workplace that would work without people, i.e. to lift a nation economy, human resources managers should be adequately positioned to provide the result. Hence the need for a result-oriented human relation in an organization is imperative. For example, when organizations train their human resources experts on how to smile at customers, such organizations should also make efforts to satisfy the desire of customers, so as not to defeat the effort of human resources managers in such organizations.

Training and development are relevant to organizations that are rapidly incorporating new technologies so as to make the current workforce more flexible and adaptable in order to enhance its chances for survival and profitability. Similarly, management development facilitates organizational continuity by preparing employees and current managers to smoothly assume higher level positions. It also helps to socialize management trainers by developing in them the right values and attitudes for working in the firm and foster organizational responsiveness by developing the skills that employee and managers need to respond faster to change organizations.

Hutchins (2009) concludes that training should be given and designed in a way that it is helpful for the trainee and even it is according to the trainee expectations and need since it is critical to the function of maintaining and development of working capabilities in employees. The organizational performance relies heavily on learning and training of employee.

Cromwell and Kolb (2004) identify two types of training given to the employees (internal and external training) the internal training is given by the company itself and external training is given by the outsiders or agencies. Chen et al. (2005) emphasis motivation of employees and persuasion as a way of enhancing their skills by training transfer process.

Egan et al (2004) observed training and development is the key for the organization success and its assessment involves taking training tests in which the test declares the trainee knowledge, attributes, skills learned during the training of the job. Hartenstein (2001) however, raved that High level of training and development orientations makes the employee more interested and more motivated in learning at workplace. Hence, Organization should use training program for the satisfaction of employees. Krishnaveni and Sripirabaa (2008). While training an employee, it should be according to the behaviour and interests of the employees because lack of diversity training especially with respect to trainee characteristics is increasingly a serious problem.

Lawal (2005), highlighted some importance of training and development which include increase in productivity, increase in workers commitment and morale, provision of opportunity for personal growth of the employees and

to Contents, page 1

Table 1: Aggregate of respondents responses

Resp.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Total
SA	68	48	30	54	60	74	23	72	40	42	85	996
A	24	44	40	30	32	15	62	20	33	46	7	353
IND	0	0	20	5	0	3	7	0	19	4	0	58
D	0	0	2	3	0	0	0	0	0	0	0	5
SD	0	0	0	0	0	0	0	0	0	0	0	0

Table 2: Relationship exist between training and development and organizational performance

	Value	Asymp. Std Error(a)	Approx Sig.
Pearson's correlation	0.564	.092	.000
Coefficient of determination	0.318	.092	.000
No valid cases	92	-	-

greater organization performance.

METHODOLOGY

In other to collect appropriate data concerning this study, primary source of data was employed using the questionnaire administered to 100 selected staff of Ikosi-Ejirin Local Government which includes Agriculture, Budget and Planning, Education, Finance, Health and Personnel departments out of the 305 staff of the council so as to have clear view of some of the responses provided by the respondent.

The questionnaire contains two parts with the first part contain 9 questions relating to the respondents personal data, while the second part contain 16 questions relating to the study and hypothesis. The questionnaire was subjected to construct and statistical validity with a critical value of 1.96 is set as the reliability coefficient.

Hypothesis: *That there is no relationship between training and development of employees and organizational productivity.*

In modeling the effect of training and development on organizational productivity, a simple linear model has been formulated. This model explains the relationship among the variables and it is stated below thus:

$P = a_0+a_IT_D+a_2S + a_3 E + a_4A$........................... (1)

Where:

P = Productivity

T_D = Training and Development

S = Job satisfaction

E = Employee Satisfaction

A= Age

The model above shows the relationship between training, development and organizational productivity. However, the relationship among the variables in terms of opinion expectation is mathematically depicted as;

$\delta p > 0 \quad \delta T_D > 0$

The above denotation implies that we expect a positive relationship between productivity, training and development.

After the collection of information, the information needs critical examination to give accurate findings. Data Analysis of this study will be done by the use of descriptive statistics such as correlation coefficient, regression were used on the specified explanatory variables.

Data analysis and interpretation

It was revealed that 50 (54.3%) of the respondents were male while 42(45.7%) of the respondents were female with an average age 32years and qualification ranging from NCE, OND, Diploma to HND. Most were at the Supervisory and managerial level of the council with majority from health, finance and personnel department. Out of the 100 questionnaires administered, 94 were received, 6 were not returned while 2 were rejected for non-clarity. (Table 1)

Hypothesis testing

Having given a clear analysis of the responses assessing the respondents through the cross tabulation of variables, the hypothesis formulated earlier are now tested. The validity of the hypothesis therefore is tested by applying correlation statistical techniques.

INTERPRETATIONS OF RESULTS

The table 2 above expounded that a moderate positive relationship exist between training and development and organizational performance with a correlation value of 0.564 (56.4%) while the coefficient of determination indicates that about 32% of variation or changes in the level of organizational performance can be attributed to training and development in an organisation.

The implication is that there is a significant relationship between training and development and organizational productivity. To this end, the explanatory with .000 and according to the decision rule, we hereby accept the alternative hypothesis which states that there is a significant relationship between training and development and organisational productivity.

This is in line with the current trend at the global scale where it is viewed that the type of human capital available in an organisation determines its level of productivity and since competition and changing business environment has compelled organisations to invest huge sum of money in the acquisition and upgrading of equipment and employed the best brain in order to be differentiated from competitors especially with the quality of its human resources.

SUMMARY OF FINDINGS

This study revealed some important findings were observed which were not limited to the case study alone but apply to many public sectors in Nigeria. They include:

a. Training and development resulted into an increase in employee's efficiency, skill and knowledge which then accounted for improved performance of the organization.

b. Training and Development surely helps employees to get to the climax of their career, like promotion and attainment of higher status.

c. Well trained employees performed better than ill-trained or untrained employees and because of the desire to excel, training and development bring out the best in them.

d. There have been improvements in the attitude of both the generality of staff and the management due to frequent exchange of idea opinions and views among them. This mostly comes up in seminar, workshop etc.

Recommendations

From the major findings of the study one realizes that there is need to make some recommendations for the local government which would help in the development of the training and development programmes.

a. The method of selecting employees for training should be defined. Employees should be sent for the most needed skill. This should be done after it has carried out the organizational analysis and strictly adhered to the outcomes to the exercise.

b. Employees should be made aware of the extent of organization spends in training and development programmes. This should not be secret to the staff if the organization intends to eliminate the impression that training period is a relaxation for the employees.

c. Management of the organization should make training a pre-requisite for promotion. This will in no small way motivate the employees as well as increase their participation in training session.

d. Performance evaluation should be objectively ascertained. These should be an improvement in the system of evaluation. The actual purposes of training is to cause a change in employee behaviour on the job and ultimately to improve the effectiveness of the organization.

e. High quality effective and efficient training and development programmes should not be sacrificial for cost.

f. High standard of professionalism of trainers, regular review of course contents, modern instructional aids etc. should be given the highest priority deserved.

REFERENCES

Brewer Ernest W.,Lim Doo Hun, Cross Marcia E. (2008) “job satisfaction and employee perception of learning in the health care management industry”, Journal of leadership studies, 1(4) .

Chen Hsin-chih, Holton E.F, Bates Reid(2005) “Development and validation of the learning transfer system inventory in Taiwan”, Human resource development quarterly 16(1).

Cormwell Susan E. ,Kolb Judith A.(2004) “An examination of work environment support factors affecting transfer of supervisory skills training to the work place”, Human resource development quarterly, 15(4).

Dale S. B. (2005): Personnel: The Management of people at work. Macmillan Publishing Co. Inc. New York

Egan Toby Marshall ,Yang Baiyin ,Bratlett Kenneth R.(2004) , The effects of organization learning culture and job satisfaction on motivation to transfer learning and turnover intention. Human resource development quarterly ;15(3)

Hartenstein Edward J. (2001) Goal orientation and practice condition as predictors of training results; Human resource development Quarterly (12) 4.

Hassan Arif (2007) “department of business administration international Islamic university Malaysia Kuala Lampur, Malaysia, Journal of European industrial training, 31(6); 435-448.

Holton Elwood .F (1999) “New employee development: A review and Re-conceptualization”

Hutchins Holly M.(2009) “In the trainer’s voice: A study of Training transfer practices”, Performance improvement quarterly ,22 (1) , (69-93).

Lawal A. A. (2005):"Developing employees for improved performance" Journal of Liberal studies vo1.No. 2, 71 – 80.

Subramony M. (2009) “A meta-analytic investigation of the relationship between HRM bundles and firm’s performance”, Human resource management, 48(5) , 745-768.

Valle Ramon, Martin Fernando, Romero Pedro M.,Dolan Shimon L (2000); Business strategy, Work process and Human resource training are they congruent? ; Journal of organizational behavior:21 ;283-297 .

Zhu, Y., Rowley, C. and Warner, M. (2009) ‘HRM in Asia’, in H. Bidgoli (ed.) *The Handbook of Technology Management*, Hoboken, NJ: Wiley.

to Contents, page 1

Author Index

G

H

I

J

K

L

M

to Contents, page 1

T

U

V

W

X

Y

Z