Mathematics, Information Technologies and Applied Sciences 2015

post-conference proceedings of extended versions of selected papers

Editors:

Šárka Hošková-Mayerová and Miroslav Hrubý

Brno, Czech Republic, 2015

ISBN: 978-80-7231-436-2

Aims and target group of the conference:

The conference **MITAV 2015** should attract in particular teachers of all types of schools and is devoted to the most recent discoveries in mathematics, informatics, and other sciences, as well as to the teaching of these branches at all kinds of schools for any age group, including e-learning and other applications of information technologies in education. The organizers wish to pay attention especially to the education in the areas that are indispensable and highly demanded in contemporary society. The goal of the conference is to create space for the presentation of results achieved in various branches of science and at the same time provide the possibility for meeting and mutual discussions among teachers from different kinds of schools and fields/specializations. We also welcome presentations by (diploma and doctoral) students and teachers who are just beginning their careers, as their views and approaches are often interesting and stimulating for other participants.

Organizers:

Union of Czech Mathematicians and Physicists, Brno branch (JČMF),
in co-operation with
Faculty of Military Technology, University of Defence in Brno,
Faculty of Science, Faculty of Education and Faculty of Economics and Administration,
Masaryk University in Brno,
Faculty of Electrical Engineering and Communication, Brno University of Technology.

Venue:

Club of the University of Defence in Brno, Šumavská 4, Brno, Czech Republic
June 18 and 19, 2015.

Scientific committee:

Prof. RNDr. Zuzana Došlá, DSc. — Czech Republic
Faculty of Science, Masaryk University, Brno

Prof. Irada Ahaievna Dzhalladova, DrSc. — Ukraine
Kyiv National Economic Vadym Getman University

Assoc. Prof. Cristina Flaut — Romania
Faculty of Mathematics and Computer Science, Ovidius University, Constanta

Assoc. Prof. PaedDr. Tomáš Lengyelfalusy, Ph.D. — Slovakia
DTI (Dubnica Technolog Institute), s.r.o.

Prof. Antonio Maturo — Italy
Faculty of Social Sciences of the University of Chieti – Pescara

Prof. Karl Hayo Siemsen — Germany
University of Applied Sciences Saarbrücken and FITT, Hochschule Emden – Leer

Programme and organizational committee:

Jaromír Baštinec	Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Mathematics
Luboš Bauer	Masaryk University in Brno, Faculty of Economics and Administration, Department of Applied Mathematics and Informatics
Jaroslav Beránek	Masaryk University in Brno, Faculty of Education, Department of Mathematics
Šárka Hošková-Mayerová	University of Defence in Brno, Faculty of Military Technology, Department of Mathematics and Physics
Miroslav Hrubý	University of Defence in Brno, Faculty of Military Technology, Department of Communication and Information Systems
Karel Lepka	Masaryk University in Brno, Faculty of Education, Department of Mathematics
Pavlína Račková	University of Defence in Brno, Faculty of Military Technology, Department of Mathematics and Physics
Jan Vondra	Masaryk University in Brno, Faculty of Science, Department of Mathematics and Statistics

Conference languages:

Czech, Slovak, English

Programme of the conference:

Thursday, June 18, 2015

12:00-14:00 Registration of the participants
14:00-14:45 Keynote lecture No. 1 (Zuzana Došlá, Czech republic)
14:45-15:15 Break
15:15-17:30 Presentations in two parallel sections

19:00-22:00 Social event

Friday, June 19, 2015

09:00-09:45 Keynote lecture No. 2 (Jiří Dan, Czech Republic)
09:45-10:15 Break
10:15-13:15 Presentations in two parallel sections
14:00 Closing

Each MITAV 2015 participant received printed collection of abstracts **MITAV 2015** with ISBN 978-80-7231-998-5. CD supplement of this printed volume contains all the accepted contributions of the conference.

Now, in autumn 2015, this **post-conference CD** was published, containing extended versions of selected MITAV 2015 contributions. The proceedings are published in English and contain extended versions of 15 selected conference papers. Published articles have been chosen from 42 conference papers and every article was reviewed by two reviewers.

Webpage of the MITAV conference:

http://mitav.unob.cz

Content:

Reviewers:

Jaromír Baštinec, Luboš Bauer, Jaroslav Beránek, František Bubeník, Jiří Jánský, Karel Lepka, Jiřina Novotná, Radovan Potůček.

Stability of the Zero Solution of Stochastic Differential Systems with Two-dimensional Brownian motion

Jaromír Baštinec, Marie Klimešová

Department of Mathematics, Faculty of Electrical Engineering and Communication Brno University of Technology,
Technická 2848/8, 61600, Brno, Czech Republic.
bastinec@feec.vutbr.cz, xklime01@stud.feec.vutbr.cz

Abstract: *The natural world is influenced by stochasticity therefore stochastic models are used to test various situations because only the stochastic model can approximate the real model. For example, the stochastic model is used in population, epidemic and genetic simulations in medicine and biology, for simulations in physical and technical sciences, for analysis in economy, financial mathematics, etc. The crucial characteristic of the stochastic model is its stability.*

This article studies the fundamental theory of the stochastic stability. There is investigated the stability of the solution of stochastic differential equations (SDEs) and systems of SDEs. The article begins with a summary of the stochastic theory. Then, there are inferred conditions for the asymptotic mean square stability of the zero solution of stochastic equation with one-dimensional Brownian motion and system with two-dimensional Brownian motion. There is used a Lyapunov function for proofs of main results.

Keywords: Brownian motion, stochastic differential equation, Lyapunov function, stochastic Lyapunov function, stability, stochastic stability.

Introduction

Stochastic modeling has come to play an important role in many branches of science and industry where more and more people have encountered stochastic differential equations. Stochastic model can be used to solve problem which evinces by accident, noise, etc. Definition of probability spaces, stochastic process, stochastic differential equation and an existence and uniqueness of solution of these equations, were mentioned in [15], [16], [17]. It was taken from B. Øksendal [13], E. Kolářová [9], B. Maslowski [11], S. Ditlevsen [3], M. Navara [12] and J. Staněk [14]. In this paper we focus on the description of the stochastic stability. Stability is studied both for difference equations and systems [5], and for differential equations and systems [1], [2], [4], [6] or [7]. The stability theory was introduced by R. Z. Khasminskii [8]. The basic principles of various types of stochastic systems are described by X.Mao [10]. In the paper we derived sufficient conditions for general system of the zero solution of the stochastic differential equation using Lyapunov function.

Definition 1 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $B_t = (B_1(t), ..., B_m(t))$ be m-dimensional Brownian motion and $b : [0, T] \times R^n \to R^n$, $\sigma : [0, T] \times R^n \to R^{n\times m}$ be measurable functions. Then the process $X_t = (X_1(t), ..., X_m(t))$, $t \in [0, T]$ is the solution of the stochastic differential equation*

$$dX_t = b(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t, \tag{1}$$

$b(t, X_t) \in R,\ \sigma(t, X_t)W_t \in R$. After the integration of equation (1) we give the solution of the SDE in the integral form

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s.$$

Assume that for every initial value $X_t(0) = X_0 \in R^n$, there exists a unique global solution which is denoted by $X(t; t_0, X_0)$. So equation (1) has the solution $X_t(0) \equiv 0$ corresponding to the initial value $X_t(0) = 0$. This solution is called the **trivial solution** or equilibrium position.

1 Stability of Stochastic Differential Equations

In 1892 A.M. Lyapunov developed a methods for determining stability without solving the equation. We are used the second Lyapunov method: Let K denote the family of all continuous nondecreasing functions $\mu : R_+ \to R_+$ such that $\mu(0) = 0$ and $\mu(r) > 0$ if $r > 0$. For $h > 0$, let $S_h = \{x \in R^n : |x| < h\}$. A continuous function $V(x, t)$ defined on $S_h \times [t_0, \infty)$ is said to be **positive-definite** (in the sense of Lyapunov) if $V(0, t) \equiv 0$ and, for some $\mu \in K$,

$$V(x, t) \geq \mu(|x|) \quad \text{for all } (x, t) \in S_h \times [t_0, \infty).$$

A function $V(x, t)$ is said to be **negative-definite** if $(-V(x, t))$ is positive-definite. A continuous non-negative function $V(x, t)$ is said to be **decrescent** (i.e. to have an arbitrarily small upper bound) if for some $\mu \in K$,

$$V(x, t) \leq \mu(|x|) \quad \text{for all } (x, t) \in S_h \times [t_0, \infty).$$

A function $V(x, t)$ defined on $R^n \times [t_0, \infty)$ is said to be **radially unbounded** if

$$\lim_{|x|\to\infty} \left(\inf_{t\geq t_0} V(x, t) \right) = \infty.$$

Let $C^{1,1}(S_h \times [t_0, \infty), R_+)$ denote the family of all continuous functions $V(x, t)$ from $S_h \times [t_0, \infty)$ to R_+ with continuous first partial derivatives with respect to every component of x and to t. Then $v(t) = V(t, X_t)$ represents a function of t with the derivative

$$\dot{v}(t) = V_t(t, X_t) + V_x(t, X_t)b(t, X_t) = \frac{\partial V}{\partial t}(t, X_t) + \sum_{i=1}^{n} \frac{\partial V}{\partial x_i}(t, X_t)b_i(t, X_t).$$

If $\dot{v}(t) \leq 0$, then $v(t)$ will not increase so the distance of X_t from the equilibrium point measured by $V(t, X_t)$ does not increase. If $\dot{v}(t) < 0$, then $v(t)$ will decrease to zero so the distance will decrease to zero, that is $X_t \to 0$.

Theorem 1 ***(Lyapunov theorem)*** *If there exists a positive-definite function* $V(x,t) \in C^{1,1}(S_h \times [t_0, \infty), R_+)$ *such that*

$$\dot{V}(x,t) := V_t(t, X_t) + V_x(t, X_t)b(t, X_t) \leq 0$$

for all $(x,t) \in S_h \times [t_0, \infty)$*, then the trivial solution is stable. If there exists a positive-definite decrescent function* $V(x,t) \in C^{1,1}(S_h \times [t_0, \infty), R_+)$ *such that* $\dot{V}(x,t)$ *is negative-definite, then trivial solution of the system is asymptotically stable.*

Suppose one would like to let the initial value be a random variable. It should also be pointed out that when $\sigma^{(x,t)} = 0$, these definitions reduce to the corresponding deterministic ones. We now extend the Lyapunov Theorem 1 to the stochastic case. Let $0 < h \leq \infty$. Denote by $C^{2,1}(S_h \times R_+, R_+)$ the family of all nonnegative functions $V(x,t)$ defined on $S_h \times R_+$ such that they are continuously twice differentiable in x and once in t. Define the differential operator L associated with equation (1) by

$$L = \frac{\partial}{\partial t} + \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(t, X_t)b_i(x,t) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}\left[\sigma(x,t)\sigma^T(x,t)\right]_{ij}.$$

The inequality $\dot{V}(x,t) \leq 0$ will be replaced by $LV(x,t) \leq 0$ in order to get the stochastic stability assertions.

Theorem 2 *If there exists a positive-definite*

(i) function $V(x,t) \in C^{2,1}(S_h \times [t_0, \infty), R_+)$ *such that* $LV(x,t) \leq 0$ *for all* $(x,t) \in S_h \times [t_0, \infty)$*, then the trivial solution of equation* (1) *is stochastically* ***stable****.*

(ii) decrescent function $V(x,t) \in C^{2,1}(S_h \times [t_0, \infty), R_+)$ *such that* $LV(x,t)$ *is negative-definite, then the trivial solution of equation* (1) *is stochastically* ***asymptotically stable****.*

(iii) decrescent radially unbounded function $V(x,t) \in C^{2,1}(R^n \times [t_0, \infty), R_+)$ *such that* $LV(x,t)$ *is negative-definite, then the trivial solution of equation* (1) *is stochastically* ***asymptotically stable in the large.***

Proof: [10], pp. 111.

2 Main results

We have a homogenous linear stochastic differential equation

$$dX_t = A(X_t)dt + GdB_t, \tag{2}$$

where $X_t = \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $G = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$, $B_t = \begin{pmatrix} B_1(t) \\ B_2(t) \end{pmatrix}$,

a, b, c, d, e, f, g are constants.

Definition 2 *Lyapunov quadratic function V is given*

$$V(X_t) = X_t^T Q\ X_t,$$

where $Q = \begin{pmatrix} p & q \\ q & p \end{pmatrix}$ *is a symmetric positive-definite matrix, i.e.* $p > 0$, $p^2 - q^2 > 0$.

Theorem 3 *Equation* (2) *is stable if*

$$LV = 2\left[aX_1^2(t) + dX_2^2(t) + (c+b)X_1(t)X_2(t) + e^2 + f^2 + g^2 + h^2\right].$$

Proof:
We compute derivation of Lyapunov function of equation (2)

$$\begin{aligned} dV(X_t) &= V(X_t + dX_t) - V(X_t) \\ &= (X_t^T + (AX_t)^T dt + (GdB_t)^T)Q(X_t + AX_t dt + GdB_t) - X_t^T Q X_t \\ &= X_t^T Q X_t + X_t^T Q A X_t dt + X_t^T Q G dB_t + (AX_t)^T dt Q X_t \\ &+ (AX_t)^T dt Q A X_t dt + (AX_t)^T dt Q G dB_t + (GdB_t)^T Q X_t \\ &+ (GdB_t)^T Q A X_t dt + (GdB_t)^T Q G dB_t - X_t^T Q X_t \\ &= X_t^T Q A X_t dt + X_t^T Q G dB_t + X_t^T A^T dt Q X_t + X_t^T A^T dt Q A X_t dt \\ &+ X_t^T A^T dt Q G dB_t + dB_t^T G^T Q X_t + dB_t^T G^T Q A X_t dt + dB_t^T G^T Q G dB_t. \end{aligned}$$

We use the rules:

$$\begin{aligned} dt \cdot dt = dt \cdot dB_1(t) = dt \cdot dB_2(t) = dB_1(t) \cdot dB_2(t) = 0, \\ dB_1(t) \cdot dB_1(t) = dB_2(t) \cdot dB_2(t) = dt. \end{aligned}$$

After modyfying we get

$$\begin{aligned} dV(X_t) &= X_t^T Q A X_t dt + X_t^T Q G dB_t + X_t^T A^T dt Q X_t + dB_t^T G^T Q X_t \\ &+ dB_t^T G^T Q G dB_t. \end{aligned}$$

In matrix form

$$\begin{aligned} dV\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} &= \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}^T \begin{pmatrix} p & q \\ q & p \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} dt \\ &+ \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}^T \begin{pmatrix} p & q \\ q & p \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix} \\ &+ \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}^T \begin{pmatrix} a & b \\ c & d \end{pmatrix}^T \begin{pmatrix} p & q \\ q & p \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} dt \\ &+ \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}^T \begin{pmatrix} e & f \\ g & h \end{pmatrix}^T \begin{pmatrix} p & q \\ q & p \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} \\ &+ \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}^T \begin{pmatrix} e & f \\ g & h \end{pmatrix}^T \begin{pmatrix} p & q \\ q & p \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}. \end{aligned}$$

We determine

$$\begin{pmatrix} e & f \\ g & h \end{pmatrix}^T \begin{pmatrix} p & q \\ q & p \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = M = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix}.$$

Then we have

$$\begin{aligned} & \left[\begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix} \right]^T \begin{pmatrix} p & q \\ q & p \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix} \\ = & \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}^T \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix} \\ = & \begin{pmatrix} m_1 dB_1(t) + m_3 dB_2(t) & m_2 dB_1(t) + m_4 dB_2(t) \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix} \\ = & m_1 dB_1(t) dB_1(t) + m_3 dB_2(t) dB_1(t) + m_2 dB_1(t) B_2(t) + m_4 dB_2(t) B_2(t) \\ = & m_1 dt + m_4 dt = tr(M) dt, \end{aligned}$$

where $tr(M)$ is trace of square matrix M.

We get

$$\begin{aligned} dV(X_t) &= 2\left[(ap+cq)X_1^2(t) + (dp+bq)X_2^2(t) + ((b+c)p \right. \\ &+ (a+d)q)X_1X_2(t) + (2q(hf+eg) + p(e^2+f^2+g^2+h^2))\left.\right] dt \\ &+ 2\left[(ep+gq)X_1(t) + (gp+eq)X_2(t)\right] dB_1(t) + 2\left[(fp+hq)X_1(t)\right. \\ &+ \left.(hp+fq)X_2(t)\right] dB_2(t). \end{aligned}$$

We apply expectation $E\{dV(X_t)\}$

$$\begin{aligned} E\{dV(X_t)\} &= 2\left[(ap+cq)X_1^2(t) + (dp+bq)X_2^2(t) + ((b+c)p \right. \\ &+ (a+d)q)X_1(t)X_2(t) + (2q(hf+eg) \\ &+ \left. p(e^2+f^2+g^2+h^2)\right] dt = LV dt. \end{aligned}$$

For $Q = I$ we get

$$LV = 2\left[aX_1^2(t) + dX_2^2(t) + (c+b)X_1(t)X_2(t) + e^2 + f^2 + g^2 + h^2\right].$$

Now we can do a discussion under which conditions the system will be stable.

The Euclidean matrix norm A on the space R^n can be define as

$$\|A\|_E := \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}^2},$$

where a_{ij} is a matrix element of the i-th line and of the j-th column of the matrix, n is number of matrix raws, m is number of matrix columns.

We denote $e^2 + f^2 + g^2 + h^2 = \|G\|^2$
and give

$$LV = 2\left[aX_1^2(t) + dX_2^2(t) + (c+b)X_1(t)X_2(t) + \|G\|^2\right]. \tag{3}$$

The Lyapunov function LV will be negative definite if and only when

$$aX_1^2(t) + dX_2^2(t) + (c+b)X_1(t)X_2(t) + \|G\|^2 \leq 0,$$

because $\|G\|^2 \geq 0$, therefore the matrix A must be sufficiently negative, to obtain a negative definite function.

Sylvester's criterion is a necessary and sufficient criterion to determine whether a matrix is positive-definite.

Theorem 4 ***(Sylvester's criterion)***
Let A be a real symetric matrix of the n-th order. For $k = 1, \ldots, n$ we denote the main subdeterminants D_k of the matrix A

$$D_k = det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{pmatrix}.$$

Then the matrix A is positive definite if and only when $D_k > 0$ pro $k = 1, \ldots, n$. And analogously the matrix A is negative definite if and only when $(-1)^k D_k > 0$ for $k = 1, \ldots, n$.

Corollary 1 *First, we consider a diagonal matrix A in the form*

$$A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}.$$

The matrix A will be negative definite under following conditions:

$$\left.\begin{array}{rcl} D_1 & = & |a_{11}| = a < 0, \\ D_2 & = & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a & 0 \\ 0 & a \end{vmatrix} = a^2 > 0 \end{array}\right\} \Rightarrow \begin{array}{l} \textit{if holds } D_1 \textit{ then the condition } D_2 \\ \textit{is obvious.} \end{array}$$

Then from (3) follows

$$aX_1^2(t) + aX_2^2(t) \quad \leq \quad -\|G\|^2$$

or

$$a\|X_t\|^2 \quad \leq \quad -\|G\|^2.$$

If the variable a is negative and also inequality $a\left\|X_t\right\|^2 \leq -\left\|G\right\|^2$ is valid, then the system is stochastically stable.

We find a solution of the stochastic system based on eigenvalues. If $a_{12} = a_{21} = 0$, then $\lambda_1 \approx a_{11}, \lambda_2 \approx a_{22} \Rightarrow \lambda_{1,2} = a$. Because a is negative we make substitution $a = -\alpha, \alpha > 0$. We give a solution of the system

$$\begin{aligned} X_1(t) &= C_1 e^{-\alpha t}, \\ X_2(t) &= C_2 t e^{-\alpha t}, \end{aligned}$$

when C_1, C_2 are constants.

Corollary 2 *We consider a diagonal matrix A in the form*

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

The matrix A will be negative definite under following conditions:

$$\begin{aligned} D_1 &= |a_{11}| = a < 0, \\ D_2 &= \begin{vmatrix} a & 0 \\ 0 & b \end{vmatrix} = ab > 0 \Rightarrow b < 0. \end{aligned}$$

Then from (3) follows

$$aX_1^2(t) + bX_2^2(t) \leq -\left\|G\right\|^2.$$

We find a solution of the stochastic system based on eigenvalues. $\lambda_1 = a$, $\lambda_2 = b$. We substitute $a = -\alpha, \alpha > 0, b = -\beta, \beta > 0$. We give a solution of the system

$$\begin{aligned} X_1(t) &= C_1 e^{-\alpha t}, \\ X_2(t) &= C_2 t e^{-\beta t}, \end{aligned}$$

C_1, C_2 are constants.

Corollary 3 *We consider a symmetric matrix A in the form*

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}.$$

The matrix A will be negative definite under following conditions:

$$\left.\begin{aligned} &D_1 = a < 0, \\ &D_2 = a^2 - b^2 > 0 \Rightarrow |a| > |b| \end{aligned}\right\} \text{ i.e. must be valid } |a| > |b| > 0.$$

Then from (3) *follows*

$$\begin{aligned} aX_1^2(t) + aX_2^2(t) + 2bX_1(t)X_2(t) &\leq -\|G\|^2 \\ a\|X(t)\|^2 + 2bX_1(t)X_2(t) &\leq -\|G\|^2 \end{aligned}$$

The variable a *must be sufficiently negative and also inequality*

$$a\|X(t)\|^2 + 2bX_1(t)X_2(t) \leq -\|G\|^2$$

must be valid, then we can say that the system is stochastically stable.

We find eigenvalues of matrix A *as the solution of the characteristic equation*

$$det(A - \lambda E) = 0,$$

where E *is the unit matrix.*

$$\begin{aligned} |A - \lambda E| = \begin{vmatrix} a-\lambda & b \\ b & a-\lambda \end{vmatrix} = (a-\lambda)^2 - b^2 &= 0, \\ (a-\lambda)^2 &= b^2, \\ |a-\lambda| &= |b|\,. \end{aligned}$$

Eigenvalues are

$$\begin{aligned} -a + \lambda_1 = |b| \Rightarrow \lambda_1 = a + |b|\,, \\ a - \lambda_2 = |b| \Rightarrow \lambda_2 = a - |b|\,. \end{aligned}$$

We substitute $a = -\alpha, \alpha > 0, |b| > 0, \alpha < |b|$ *, i.e.*

$$\begin{aligned} \lambda_1 &= -\alpha + |b|\,, \\ \lambda_2 &= -\alpha - |b|\,. \end{aligned}$$

For the eigenvalue $\lambda_1 = -\alpha + |b|$ *we find the eigenvector*

$$v_1 = (v_{11}, v_{12}).$$

There is any nonzero vector which fulfills a following relation

$$\begin{aligned} (A - \lambda_1 E)\, v_1 &= 0 \\ \begin{pmatrix} a - (a+|b|) & b \\ b & a - (a+|b|) \end{pmatrix} v_1 &= 0 \end{aligned}$$

For $b > 0$ we choose an arbitrary vector $v_1 = (1,1)^T$, for $b < 0$ we choose $v_1 = (-1,1)^T$.

Then

$$\begin{array}{lll} \textit{for} & b > 0 & \textit{is} \;\; X_1(t) = (1,1)^T e^{(-\alpha+b)t} \\ \textit{for} & b < 0 & \textit{is} \;\; X_1(t) = (-1,1)^T e^{(-\alpha+b)t} \end{array}$$

For the eigenvalue $\lambda_1 = -\alpha - |b|$ we find an eigenvector

$$v_2 = (v_{21}, v_{22})$$

$$\begin{array}{c} (A - \lambda_1 E)\; v_2 = 0 \\ \left(\begin{array}{cc} a - (a - |b|) & b \\ b & a - (a - |b|) \end{array} \right) v_2 = 0 \end{array}$$

For $b > 0$ we choose an arbitrary vector $v_2 = (1,-1)^T$, for $b < 0$ we choose $v_2 = (1,1)^T$.

Then

$$\begin{array}{lll} \textit{for} & b < 0 & \textit{is} \;\; X_2(t) = (1,1)^T e^{(-\alpha-b)t} \\ \textit{for} & b > 0 & \textit{is} \;\; X_2(t) = (1,-1)^T e^{(-\alpha-b)t} \end{array}$$

The general solution is given by a linear combination $X_t = C_1 X_1(t) + C_2 X_2(t)$, with arbitrary constants C_1, C_2.

Corollary 4 *We consider a symmetric matrix A in the form*

$$A = \left(\begin{array}{ccc} a & 0 & b \\ 0 & a & 0 \\ b & 0 & a \end{array} \right).$$

The matrix A will be negative definite under following conditions:

$$\left. \begin{array}{l} D_1 = a < 0, \\ D_2 = a^2 > 0, \;\; D_2 \textit{ follows from } D_1, \\ D_3 = a^3 - ab^2 < 0 \Rightarrow a(a^2 - b^2) < 0 \Leftrightarrow a < 0 \wedge a^2 > b^2, \end{array} \right\} \Rightarrow |a| > |b|\,.$$

We find eigenvalues of matrix A as the solution of the characteristic equation

$$\left| \begin{array}{ccc} a - \lambda & 0 & b \\ 0 & a - \lambda & 0 \\ b & 0 & a - \lambda \end{array} \right| = 0,$$

$$\begin{aligned}(a-\lambda)^3-(a-\lambda)b^2 &= 0,\\ (a-\lambda)((a-\lambda)^2-b^2) &= 0 \Leftrightarrow (a-\lambda)=0 \vee (a-\lambda)^2-b^2=0,\end{aligned}$$

$$\begin{aligned}\lambda_1 = 0 \Rightarrow X_1(t) = e^0 = 1,\\ \lambda^2 - 2a\lambda + (a^2-b^2) = 0,\end{aligned}$$

$$\lambda_{2,3} = \frac{2a \pm \sqrt{4a^2-4(a^2-b^2)}}{2} \Rightarrow \lambda_{2,3} = a \pm |b|\,.$$

We substitute $a = -\alpha, \alpha > 0, |b| > 0, \alpha > |b|$ *, i.e.*

$$\begin{aligned}\lambda_2 &= -\alpha + |b|\,,\\ \lambda_3 &= -\alpha - |b|\,.\end{aligned}$$

For the eigenvalue $\lambda_2 = -\alpha + |b|$ *we find the eigenvector*

$$v_2 = (v_{21}, v_{22}, v_{23}).$$

There is any nonzero vector which fulfills a following relation

$$\begin{pmatrix} a-(a+|b|) & 0 & b \\ 0 & a-(a+|b|) & 0 \\ b & 0 & a-(a+|b|) \end{pmatrix} (A-\lambda_2 E)\, v_2 = 0 \; v_2 = 0$$

For $b > 0$ *we choose an arbitrary vector* $v_2 = (1,0,1)^T$*, for* $b < 0$ *we choose* $v_2 = (1,0,-1)^T$.

Then

$$\begin{aligned}&\textit{for} \quad b > 0 \quad \textit{is} \quad X_2(t) = (1,0,1)^T e^{(-\alpha+b)t},\\ &\textit{for} \quad b < 0 \quad \textit{is} \quad X_2(t) = (1,0,-1)^T e^{(-\alpha+b)t}.\end{aligned}$$

For the eigenvalue $\lambda_3 = -\alpha - |b|$ *we find an eigenvector*

$$v_3 = (v_{31}, v_{32}, v_{33}),$$

$$\begin{pmatrix} a-(a-|b|) & 0 & b \\ 0 & a-(a-|b|) & 0 \\ b & 0 & a-(a-|b|) \end{pmatrix} (A-\lambda_3 E)\, v_3 = 0, \; v_3 = 0.$$

For $b > 0$ we choose an arbitrary vector $v_3 = (1, 0, -1)^T$, for $b < 0$ we choose $v_3 = (1, 0, 1)^T$.

Then

$$\begin{aligned} &\textit{for} \quad b < 0 \quad \textit{is} \quad X_3(t) = (1, 0, 1)^T e^{(-\alpha - b)t}, \\ &\textit{for} \quad b > 0 \quad \textit{is} \quad X_3(t) = (1, 0, -1)^T e^{(-\alpha - b)t}. \end{aligned}$$

The general solution is given by a linear combination $X_t = C_1 X_1(t) + C_2 X_2(t) + C_3 X_3(t)$, with arbitrary constants C_1, C_2, C_3,

$$\begin{aligned} &\textit{for} \quad b > 0 \quad \textit{is} \quad X_t = C_1 + C_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} e^{(-\alpha + b)t} + C_3 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} e^{(-\alpha - b)t}, \\ &\textit{for} \quad b < 0 \quad \textit{is} \quad X_t = C_1 + C_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} e^{(-\alpha + b)t} + C_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} e^{(-\alpha - b)t}. \end{aligned}$$

***Note:** It is a solution of differential equation without a stochastic element. We have demonstrated the matrix A must be dominant for the stability of the system,*

$$\|A\| \gg \|G\| .$$

2.1 Examples

Example 1 *We have stochastic differential equation in the form*

$$d \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} dt + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}. \tag{4}$$

We determine stability of solution using derivation of Lyapunov function

$$\begin{aligned} dV \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} &= 2X_1(t)dB_1(t) + 2X_2(t)dB_2(t) + 4dt, \\ E\left\{ dV \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} \right\} &= 4dt = LV dt. \end{aligned}$$

Function $LV = 4 > 0$ is positive-definite. Trivial solution of system (4) is unstable.

Example 2 *We have stochastic differential equation in the form*

$$d \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} dt + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}. \tag{5}$$

We determine stability of solution

$$\begin{aligned} dV\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} &= 2(-2X_1(t)^2(t) - 2X_2(t)^2(t) - 2X_1(t)X_2(t) + 2)dt \\ &+ 2X_1(t)dB_1(t) + 2X_2(t)dB_2(t), \\ E\left\{dV\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}\right\} &= 2(-2X_1^2(t) - 2X_2^2(t) - 2X_1(t)X_2(t) + 2)dt = LVdt. \end{aligned}$$

Function is negative-definite for $LV < 0$, i.e.

$$\begin{aligned} 2(-2X_1^2(t) - 2X_2^2(t) - 2X_1(t)X_2(t) + 2) &< 0, \\ |X_1(t) + X_2(t)| &> \sqrt{1 - X_1(t)X_2(t)}, \end{aligned}$$

for $X_1(t)X_2(t) \leq 1$, then trivial solution of system (5) is stable.

3 Conclusion

In this paper it was defined stability and stochastic stability of the stochastic differential equations. It was computed specific examples by using Lyapunov theorem. Such type of equations can be used also in biomedical engineering, in meteorology, epidemic modeling, predicting economics, etc.

Acknowledgement

This research was supported by Grant FEKT-S-11-2-921 of Faculty of Electrical Engineering and Communication, BUT.

Reference

[1] BAŠTINEC, J.; DZHALLADOVA, I.:*Sufficient conditions for stability of solutions of systems of nonlinear differential equations with right-hand side depending on Markov's process.* In 7. konference o matematice a fyzice na vysokých školách technických s mezinárodní účastí. 2011. p. 23 - 29. ISBN 978-80-7231-815-5.

[2] DIBLÍK, J., KHUSAINOV, D.Y., BAŠTINEC, J., RYVOLOVÁ, A.: *Exponential stability and estimation of solutions of linear differential systems with constant delay of neutral type.* In 6. konference o matematice a fyzice na vysokých školách technických s mezinárodní účastí. Brno, UNOB Brno. 2009. p. 139 - 146. ISBN 978-80-7231-667-0.

[3] DITLEVSEN, S., BATZEL, J., BACHAR, M.: *Stochastic biomathematical models*, Heidelberg: Springer, 2013, 206 p.

[4] DURRETT, R.: *Probability: theory and examples*, 3. ed. Belmont, CA: Thomson Brooks/Cole, 2005, 497 s. ISBN 05-344-2441-4.

[5] IGNATYEV, A. O; IGNATYEV, O. QUADRATIC FORMS AS LYAPUNOV FUNCTIONS IN THE STUDY OF STABILITY OF SOLUTIONS TO DIFFERENCE EQUATIONS. *Electronic Journal of Differential Equations.* 2011, (19): 1â“21. ISSN 1072-6691. Available from: http://ejde.math.txstate.edu

[6] DZHALLADOVA, I.A.: *Optimization of stochastic systems*, Kiev, KNEU Press, 2005. ISBN 966-574-774-6.

[7] DZHALLADOVA, I.; BAŠTINEC, J.; DIBLÍK, J.; KHUSAINOV, D.: *Estimates of exponential stability for solutions of stochastic control systems with delay.* Abstract and Applied Analysis. 2011. 2011(1). p. 1 - 14. ISSN 1085-3375. (IF=1,318).

[8] KHASMINSKII, R.: *Stochastic stability of differential equations.* New York: Springer Berlin Heidelberg, 2011, 358 s. ISBN 978-3-642-23279-4.

[9] KOLÁŘOVÁ, E.: *Stochastické diferenciální rovnice v elektrotechnice.* Thesis. Brno: BUT, 2006, 26 s. ISBN 80-214-3330-2.

[10] MAO, X.: *Stochastic differential equations and applications.* Chichester: Horwood Pub., 2008, 422 s. ISBN 978-1-904275-34-3.

[11] MASLOWSKI, B., MILOTA, J.: *Proceedings of Seminar in Differential Equations.* Plzeň: Vydavatelský servis, 2006, 118 s. ISBN 80-86843-14-9.

[12] NAVARA, M.: *Pravděpodobnost a matematická statistika.* Skriptum. Praha: FEL ČBUT, 2007, 240 s. ISBN 978-80-01-03795-9.

[13] ØKSENDAL, B.: Stochastic Differential Equations. *An Introduction with Applications*, Springer-Verlag, 1995.

[14] STANĚK, J.: *Stochastické diferenciální rovnice*, KDM - MFF UK, 2011.

[15] KLIMEŠOVÁ, M.: *Stochastic Differential Equations*, Student EEICT. Brno: LITERA, 2014. s. 150-154. ISBN: 978-80-214-4924-4.

[16] KLIMEŠOVÁ, M.: *Stability of the Stochastic Differential Equations*, Student EEICT. Brno: BUT, 2015. s. 526-530. ISBN: 978-80-214-214-5148-3.

[17] KLIMEŠOVÁ, M.; BAŠTINEC, J. *Application of Stochastic Differential Equations.* MITAV. Brno: UO, 2014. s. 1-6. ISBN: 978-80-7231-961-9.

ITERATIVE ROOTS OF SET TRANSFORMATIONS AND THEIR USE

Jaroslav Beránek
Faculty of Education, Masaryk University
Poříčí 7, 603 00 Brno, Czech Republic
beranek@ped.muni.cz

Abstract: *The article is devoted to the problem of existence and construction of iterative roots of mappings of the sets into themselves and their possible use while solving tasks. The main part of the article is devoted to the application of the stated theory. First, the necessary and sufficient condition for the existence of iterative roots of all orders is given, further the condition is specified for the existence of the iterative root of order two.*

Key words: Mapping, orbit, cycle, iteration, iterative root

INTRODUCTION

The article is devoted to the iterative roots of set transformations and their use while solving tasks. Althought at first sight this topic seems to be considerably distant from the university mathematics teaching, the reverse is true. The substance of the existence and construction of iterative roots lies in the approach to mappings and functions from the discrete point of view, when we understand them as monounary algebras and represent them with the help of the vertex graphs. Such an approach enables effective solving of many problems and tasks from different mathematics areas. Moreover, in comparison with the classical approach to mappings and functions from the continuous perspective, it contributes to the deeper insight to its mathematical essence. The discrete interpretation of functions appears only seldom at the mathematics teaching at high schools and universities, although it can be extremely beneficial for participants in higher levels of the Mathematical Olympiad (see [4], [6], [12]). The considerations, which are used while formulating definitions, theorems and proofs in the iteration theory, especially the ones used when solving the problems of existence and construction of iterative roots of functions on finite sets, can be used as the suitable topic for students´ individual scientific activity while their mathematical abilities development. Students can thus discover their own numerous nontrivial results without studying formally complicated theories, too distant from the commonly discussed topics in the regular lessons. Now, let us remind some necessary terms and theorems from the functions iterative theory.

1. ITERATIONS OF SET TRANSFORMATIONS, VERTEX GRAPHS

The mapping $f: X \to X$ of the set X into itself will be called the transformation of the set X. For $n \in \mathbf{N_0}$ let us define the n-th iteration f of the set X as follows:
$f^0(x) = x$, $f^1(x) = f(x)$, $f^n(x) = (f \circ f^{n-1})(x)$ for every $x \in X$; in the shortened form $f^n = f \circ f^{n-1}$.

If the transformation f is a bijective mapping of the set X into itself, the definition of the given set iterations can be broadened also for a non-negative integer n in the following way: let us denote f^{-1} as an inverse function to the function f on the set X, then $f^{-2} = f^{-1} \circ f^{-1}$, $f^{-n} = (f^{-1})^n$. It is necessary to distinguish between the notation of the n-th iteration of the function f, which is f^n (the value of the iteration for the element x is $f^n(x)$), and the expression $[f(x)]^n$.

Every transformation f of the set X determines the equivalence $\sim_f$ on X as follows: $x \sim_f y$, if and only if there exists a pair of positive integers m, n that $f^m(x) = f^n(y)$. The blocks of the decomposition of the set X determined by the equivalence $\sim_f$ are called orbits of the transformation f, in short f-orbits. The set containing elements $x, f(x), f^2(x), f^3(x), \ldots$ is called the iterative sequence starting in x or also the f-splinter of the element x.

Let k be a natural number, then the cycle of the order k (k-cycle) of the mapping $f: X \to X$ is the set $\{x_0, x_1, \ldots, x_{k-1}\}$ of the set X elements for which there applies $f(x_m) = x_{m+1}$ for $0 \leq m < k-1$ and $f(x_{k-1}) = x_0$. The orbit containing a cycle is called the cyclic one, otherwise the acyclic one. For $k = 1$, the element $x \in X$ with the property $f(x) = x$ is called the fixed point of the transformation f. For cyclic orbits, there is an important term of the depth of the element x (below the cycle) which is denoted $h(x)$ and defined as the least non-negative number for which $f^{h(x)}(x)$ is the element of the cycle. All elements of the cycle are of the depth 0.

Let us give some orbit properties which will be further used (see [14]):

- Every orbit contains at most one cycle.
- The orbit is acyclic if and only if for its every element there applies that the corresponding iterative sequence contains infinitely many elements.
- Every finite orbit is cyclic (the chain ending in the cycle is not infinite, although it contains infinitely many elements).

In the case of the injective transformation f, the orbits are isolated cycles, two-sidedly infinite chains, or infinite chains bounded from below by the least elements; if f is a bijection, its orbits are either cycles or two-sidedly infinite chains. The set of orbites of the function f is also called the orbit structure. The graphic representation of the orbits is the vertex graph.

Here follow illustrative examples.

a) Let $X = \{1,2,3,\ldots,8\}$, the transformation f is defined: $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 4 & 4 & 6 & 8 & 8 & 8 & 8 \end{pmatrix}$.

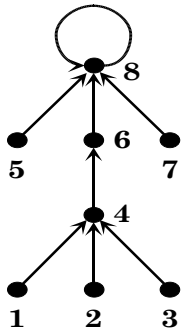


Fig.1. Vertex graph of transformation f defined by the matrix.

b) Let $X = \boldsymbol{R}$, for every $x \in X$ there applies $f(x) = -x$. The only fixed point is number zero, for other elements of the set X there applies $f^2(x) = x$. The orbits of the function f are then one loop and uncountably many cycles of order 2.

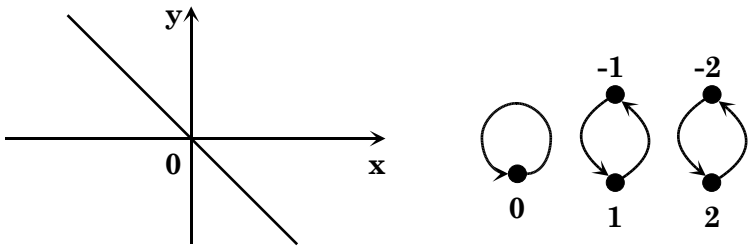


Fig. 2. Vertex graph of transformation $f(x) = -x$.

c) Let $X = \boldsymbol{R} - \{0\}$, for every $x \in X$ there applies $f(x) = x^{-1}$. For $x \in \{-1,1\}$ there applies $f(x) = x$, for other elements of the set X there applies $f^2(x) = x$. The orbits of the function f are then two loops (fixed points) and uncountably many cycles of order 2.

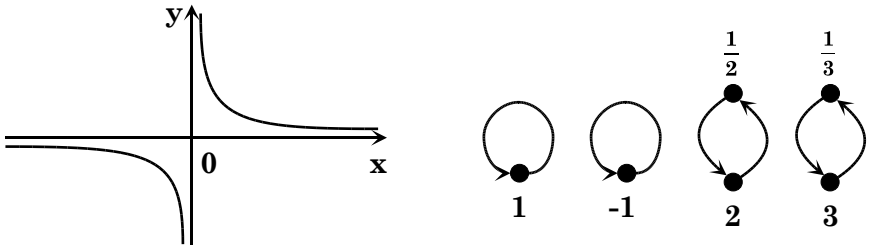


Fig. 3. Vertex graph of the transformation $f(x) = x^{-1}$.

2. ITERATIVE ROOTS

Let $X \neq \phi$, let f be the mapping of the set X into itself, the number $m \in \boldsymbol{N}$, $m > 1$. The main problem of the iterative theory is to find such an arbitrary mapping g of the set X into itself that for every element x of the set X there applies:

$$g^m = f$$

The mapping g is called the iterative root of the order m of the function f or the m-th iterative root of the function f. Let us illustrate the term of the second iterative root in Fig. 4, where there are vertex graphs of transformations f, g of the set $X = \{1,2,3,4,5,6,7,8\}$.

$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 6 & 6 & 8 & 8 & 8 & 8 & 8 \end{pmatrix}$, $g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 4 & 4 & 6 & 8 & 8 & 8 & 8 \end{pmatrix}$, there applies $g^2 = f$.

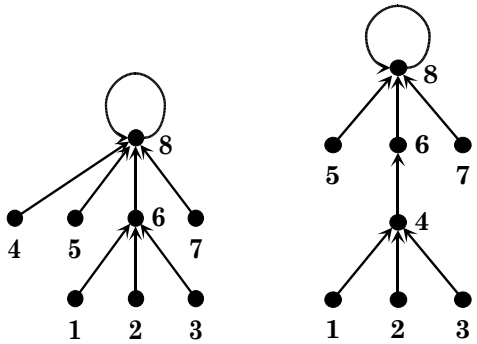


Fig. 4: The mapping g is the second iterative root of the mapping f.

Now let us briefly outline the general theory of the existence and construction of the iterative roots. For the didactic purposes, it is fundamental that in special cases (real elementary functions, bijective functions, …) it is not necessary to apply complicated theorems of the general theory while solving functional equations of one variable, but there exists a more efficient solution. The following twelve theorems are taken from the publication [14], where you can find their proofs and other details.

Theorem 1: Let $X \neq \phi$, let f, g be such mappings of the set X that $g^m = f$, $m \in \mathbf{N}$. Then the mapping g is surjective if and only if f is surjective.

Theorem 2: Let $X \neq \phi$, let f, g be such mappings of the set X that $g^m = f$, $m \in \mathbf{N}$. Then the mapping g is injective if and only if f is injective.

Theorem 3: Let $X \neq \phi$, let f, g be such mappings of the set X that $g^m = f$, $m \in \mathbf{N}$. Then the mapping g is bijective if and only if f is bijective.

Theorem 4: Let g be the m-th iterative root ($m \in \mathbf{N}$, $m \geq 2$) of the mapping f of the non-empty set X. Then every g-orbit is the union of p f-orbits, where p/m. If $p < m$, then g-orbits are n-cyclic, and p/n. In addition, all f-orbits are $\frac{n}{p}$-cyclic and at the same time the greatest common divisor (GCD) of numbers m, n equals p.

Theorem 4 describes properties of iterative roots provided that they exist. Now let us state the general necessary and sufficient conditions for the existence of iterative roots.

Definition: Let f be the mapping of the set X into itself, let r, m be natural numbers with the property r/m. Let the mapping f contain at least r orbits and let there be given r f-orbits. These orbits will be denoted m-mateable (by any mapping g), if g is the m-th iterative root of the function f, if it has one orbit and represents the union of the given r f-orbits into themselves. For $r = 1$ this only f-orbit is called m-self- mateable.

Theorem 5: If in the previous definition there applies $r < m$, then the necessary condition for the m-mateability of r f-orbits is the fact that each of them is k-cyclic (with the same k) and there applies that GCD $(k, \frac{m}{r}) = 1$. The corollary of this theorem is, among others, the fact that the acyclic f-orbit cannot be m-self- mateable for any m.

Theorem 6: An arbitrary mapping of a non-empty set has the m-th iterative root ($m \in \mathbf{N}$) if and only if the set of orbits of this mapping can be decomposed to disjoint blocks with following properties:

1° The number of orbits in each block is finite and it is the divisor of the number m.

2° Orbits in each block are m-mateable.

Theorem 7: For the existence of the m-th iterative root ($m \in \mathbf{N}$, $m \geq 2$) of the mapping $f: X \to X$ it is sufficient if in the orbit structure of the function f there exist for each occurring orbit type either infinitely many orbits of such type or their number is divisible by the number m.

Theorem 8: Let f be the bijection of any set into itself. Let us denote l_0 the number of the two-sidedly infinite chains, l_k the number of the k-cycles of the mapping f, $k \in \mathbf{N}$. Then there exists the m-th iterative root ($m \geq 2$, $m \in \mathbf{N}$) of the mapping f if and only if for every non-

negative number k there applies either $l_k = \infty$ or $d_k | l_k$, where $d_0 = m$, $d_k = \frac{m}{m_k} (k \in \boldsymbol{N})$, and m_k denotes the greatest common divisor of the number m, which is coprime with the number k.

Theorem 9: Let $f: X \to X$ be the bijection such that for every $k \in \boldsymbol{N}_0$ there applies either $l_k = 0$ or $l_k = \infty$ (according to the notation of Theorem 8). Then f has the m-th iterative root for every natural number m. For the orbits of this iterative root there also applies either $l_k = 0$ or $l_k = \infty$ for all $k \in \boldsymbol{N}_0$.

Theorem 10: Every strictly increasing and continuous bijection $\boldsymbol{R}$ on $\boldsymbol{R}$ has iterative roots of all orders.

Theorem 11: The strictly decreasing and continuous bijection of the set $\boldsymbol{R}$ has iterative roots of all orders if and only if it has either infinitely many 2-cycles or none.

Theorem 12: Every strictly decreasing and continuous bijection $\boldsymbol{R}$ has iterative roots of all odd orders.

3. USE OF ITERATIVE THEORY – EXAMPLES

All following poroblems are taken from publications [4] and [6].

Problem 1: At the 28th International Mathematical Olympiad in 1987 in Havana there was set the following task:

Prove that there is no function f from the set of non-negative integers ($\boldsymbol{N}_0 = \{0, 1, 2,...\}$) into itself such that $f(f(n)) = n + 1987$ for every $n \in \boldsymbol{N}_0$..

Let us use the iterative theory. The function $\varphi(x) = x + 1987$ is not a bijection on the set $\boldsymbol{N}_0$, but it is injective. It does not have fixed points, its orbits are mutually isomorphic chains bounded from below. There are *1987* chains, their least elements are *0, 1, ..., 1986*. The vertex graph is outlined in Fig. 5:

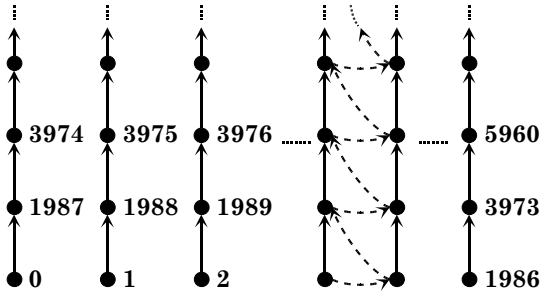


Fig. 5. Vertex graph of function $\varphi(x) = x + 1987$.

The difficulty of this task is to prove that the function φ does not have the second iterative root, i.e. that the φ-orbits are not *2*-mateable. The main idea of this proof is the fact that there is an odd number of orbits. The orbits are not cyclic, therefore, according to Theorem 5, they cannot by self-mateable. According to Theorems 5 and 6, for the existence of the second iterative root the number of φ-orbits have to be even (orbits can be mated only in pairs). This is not true, so the function φ does not have the iterative root of order *2*.

Note: With the help of iterative theory it is possible to generalize the Problem 1. The first question is if the function $\varphi(n) = n + 1987$ has any own iterative roots (of the order greater

than 1). With respect to Theorems 5 and 6 it is obvious that we are searching the possibility of the mating of the existing *1987* orbits. As *1987* is a prime number, the only possible own iterative root is the root of order *1987*. Then there exists the function $f: \mathbf{N}_0 \to \mathbf{N}_0$ with the property $f^{1987}(n) = n + 1987$, $n \in \mathbf{N}_0$. This function f is the successor function ν_0 on $\mathbf{N}_0$, defined by the formula $f(n) = n + 1$. The next question is to find out if there are own iterative roots of the function $\varphi(n) = n + c$, $n \in \mathbf{N}_0$, $c \in \mathbf{N}$. The vertex graph now contains just c isomorphic orbits (chains bounded from below with the least elements $0, 1, \ldots, c - 1$). These chains have to be mated. Similarly as above, there always exists the iterative root of the order c (which is the function $f(n) = n + 1$). Further, there always exist iterative roots of these orders which are the dividers of the number c. Therefore, the iterative root of the order 2 exists if the number c is even. If the Problem 1 were set for the function $f(n) = n + 1988$, the second iterative root would exist (further there would exist iterative roots of orders *4, 7, 14, 28, 71, 142, 284, 497, 994, 1988*). Let us illustrate the whole situation for $c = 2$.

Problem 2: Prove that there exist just two functions in $\mathbf{N}_0$ which satisfy the formula

$$f^2(n) = n + 2.$$

The vertex graph of the function $\varphi(n) = n + 2$ contains just two orbits (the chains of even and odd non-negative integers). From the general theory there follows that the only possible own iterative root is the one of the order *2*. The orbits are not *2*-self- mateable, so it is necessary to mate them together. The decomposition of the orbit set to the blocks by two orbits is the only one possible; the order of the orbits is important while mating them, so there are just two possibilities of the mating. Both functions are represented by the following formulas and shown in Fig. 6.

$$f_1(n) = n + 1,\ n \in \mathbf{N}_0 \qquad f_2(n) = \begin{cases} n-1 & \text{for } n \text{ odd}, \\ n+3 & \text{for } n \text{ even}. \end{cases},\ n \in \mathbf{N}_0.$$

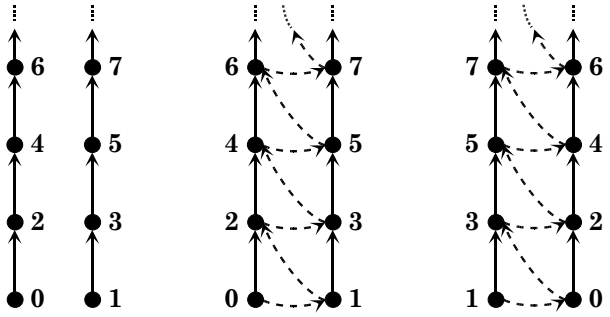


Fig 6. Solution of Problem 2.

Problem 3: Problem from the 20th International Mathematical Olympiad in Romania in 1978.

" Prove that there exist the function $f: \mathbf{N} \to \mathbf{N}$ satisfying the equation $f(f(n)) = n^2$."

First, let us give the author solution without a commentary. Let there be the sequence n_k, $k = 1, 2, \ldots$, where $n_1 = 2, n_2 = 3, n_3 = 5, \ldots$, which contains all natural numbers which are not the squares of an integer, in the natural ordering. Let us set $n_{k,m} = (n_k)^{2^m}$ for $k \in \mathbf{N}$, $m \in \mathbf{N}$. Then there holds $n_{k,m+1} = (n_k)^{2^{m+1}} = [(n_k)^{2^m}]^2 = (n_{k,m})^2$, and for any $n \geq 2$ there exists the only pair

of numbers k, m with the property $n = n_{k,m}$. Let us now define $f(n)$ as follows: $f(1) = 1$, for k odd $f(n_{k,m}) = n_{k+1,m}$, for k even $f(n_{k,m}) = n_{k-1,m+1}$. Then there applies $f^2(n) = n^2$.

The solution with the help of the iteration theory. Fig. 7 illustrates the vertex graph of the function $q(n) = n^2$ in the set $\boldsymbol{N}$. The vertex graph contains one isolated fixed point $n = 1$ and countably many infinite chains bounded from below. For every $m \in \boldsymbol{N}$, $m \geq 2$ there holds $f^m(1) = 1$, so the fixed point $x = 1$ is always m-self-mateable. Further, the set of chains can be decomposed into blocks by two, in each block the chains are 2- mateable. Therefore, the second iterative root of the function $q(n) = n^2$ in $\boldsymbol{N}$ does not exist.

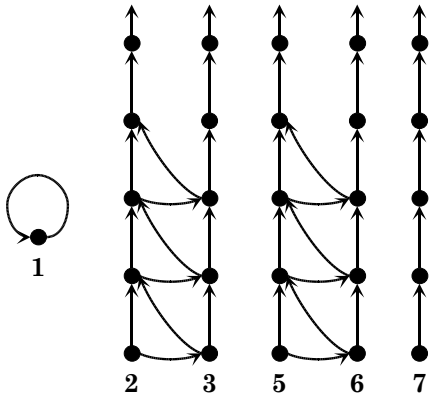


Fig. 7. Sollution of Problem 4.

Properly speaking, the author solution is only a formal mathematical description of the solution with the help of the vertex graph (by mating the chains).

Problem 4: Problem from the Mathematical Olympiad Correspondence Seminar 1983/84.

"Let f, g be mappings of the set A into itself. Let us call the function f as the n-th functional root g ($n \in \boldsymbol{N}$), if $f^n(x) = g(x)$ for every $x \in A$. Let us define $f^1(x) = f(x)$, $f^{n+1}(x) = f[f^n(x)]$.

a) Prove that the function g mapping the set $\boldsymbol{R}^+$ into itself and defined by the formula $g(x) = \frac{1}{x}$ has infinitely many n-th functional roots for every $n \geq 2$.

b) Prove that there exists the injective mapping $\boldsymbol{R}$ into $\boldsymbol{R}$, which does not have the n-th functional root for any $n \geq 2$."

For the sake of authenticity, the wording of the task is in the original version, although now the term functional root is replaced by the term iterative root.

a) The vertex graph of the function $g(x) = \frac{1}{x}$ in $\boldsymbol{R}^+$ contains one fixed point $x = 0$ and uncountably many 2-cycles. For every $n \in \boldsymbol{N}$, $n \geq 2$ they can be decomposed into blocks by n 2-cycles, in every block the 2-cycles are mutually n- mateable. As the fixed point (in the orbital structure represented by the loop) is self- mateable for any n, there follows the existence of the n-th iterative root for every n. The above described decomposition of the set of 2-cycles into blocks can be performed for any n in uncountably many ways, so for every $n \in \boldsymbol{N}$, $n \geq 2$ there exist uncountably many n-th iterative roots.

b) The vertex graph of the bijection contains only cyles and two-sidedly infinite chains. From the general theory there follows that for the bijective mapping not to have any own iterative roots there suffices if its set of orbits contains just one two-sidedly infinite chain. Let us consider the function $f(x) = x + 1$ for every $x \in \boldsymbol{Z}$, $f(x) = x$ for $x \in \boldsymbol{R} - \boldsymbol{Z}$. This mapping f contains the only one two-sidedly infinite chain and infinitely many loops.

4. ITERATIVE ROOTS OF FINITE SETS TRANSFORMATIONS

Further, we will consider the question of the existence and construction of iterative roots of the transformations of finite sets. We will show that in the case of finite sets it is not necessary to apply the general theory (Theorems 1 – 12), but it is possible to proceed in a different way. In the next text the finite set will be denoted as X. The monoid of all transformations of the set X wil be denoted as $T(X)$, the symmetric group (the group of permutations) of the set X will be denoted as $G(X)$.

Let us first give the characterization of finite sets which have iterative roots of all orders. It can be easily proved (see e.g. [9]) that identity is the only permutation of finite sets with such a property. The situation is more difficult for non-bijective mappings. Let us remind that $h(x)$ denotes the depth of the element x in the given f-orbit.

Theorem 13: (see [9]) Let X be a finite set. If the transformation f of the set X has an iterative root of the order $m = GCD\{1,2,...,card\ X\}$, then there applies $f^2 = f$ (and therefore f is its r-th iterative root for every $r \in \boldsymbol{N}$).

Proof: First, let us show that for every transformation $g \in T(X)$ there holds $g^{2m} = g^m$. As for any $x \in X$ there holds $h(x) \leq card\ X - 1$, the element $g^m(x)$ belongs to the cycle of the mapping g for $m = NSN\{1, 2,..., card\ X\}$. The order of an arbitrary cycle is at most equal $card\ X$, m is the multiple of the order of all cycles, so for every $x \in X$ lying in some of the cycles of the transformation g there holds $g^m(x) = x$. Further there follows $g^{2m}(x) = g^m[g^m(x)] = g^m(x)$ for any $x \in X$, so for every $g \in T(X)$ there holds $g^{2m} = g^m$. Based on the premise about the existence of the m-th iterative root of the transformation g there holds $f^2 = (g^m)^2 = g^{2m} = g^m = f$.

Corollary: (see [9]) The transformation f of the finite set X has iterative roots of all orders if and only if $f^2 = f$. Then it is its own iterative root of an arbitrary order. (The example of the vertex graph of the transformation with the given property is in Fig. 8)

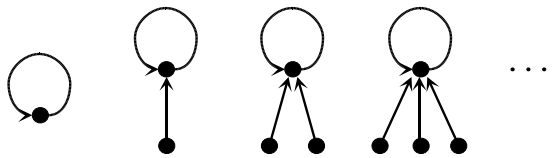


Fig. 8: Example of orbit types of set transormations f with property $f^2 = f$.

Proof: Let $f^2 = f$; then $f^r = f$ for every $r \in \boldsymbol{N}$, so f is its own iterative root of any order. On the contrary, let f have iterative roots of all orders. Then specially it has the root of the order $m = GCD\{1, 2,...,card\ X\}$ and from the previous Theorem there folows $f^2 = f$.

In the next part of the article we will limit ourselves only to the question of the second iterative roots of finite mappings. Such a restriction to the only order (*m=2*) enables the far more accurate characterization of the appropriate final mappings. Let us remind that this problem was dealt with mainly by M. Snowdon and J. Howie in [13]. Because of the didactic character of the article and its extent, we will not give proofs to the theorems stated further. All of them can be found in the above mentioned article [13]. Let us introduce the following denotation. Let *X* be a finite set, *f* be a transformation on *X*. Then π_f is an equivalence relation on *X* corresponding to *f*, so there applies: $(x,y) \in \pi_f \Leftrightarrow f(x) = f(y)$. The set *G(X)* with the operation of the mapping composition is a symmetric group of the permutations of the set *X*. If $f \in G(X)$, then also $g:X \to X$ with the property $g^2 = f$ has to belong to *G(X)* (see [14]). The question of the existence of the second iterative roots in *G(X)* can be solved separately. Let us remind that the orbit structure of each permutation $f \in G(X)$ contains only cycles. For any transformation $f \in T(X)$ of the finite set *X* then holds that each its orbit is cyclic.

Theorem 14: Let *X* be a finite set. The element $f \in G(X)$ is the second iteration of some permutation if and only if for every even number *k* the orbit structure of the permutation *f* contains the even number of *k*-cycles.

Definition: Let *f* be the transformation of the finite set *X*. Let *k* be the least non-negative integer with the property $f^k(X) = f^{k+1}(X) = \ldots$. Then this number is called the contract coefficient of the transformation *f* and is denoted as *cont f*. The subset $f^k(X) \subset X$ is called the stable range and is denoted as *stran f*. It is evident that *cont f* equals the maximum depth of the element below the cycle in the orbit structure of the transformation *f*. *Stran f* is then the union of the cycles of the transformation *f*, i.e. $f|stran\ f$ is the permutation of *stran f*.

Definition: We will say that $f \in T(X)$ is the quasi-quadratic element in *T(X)* (or shortly the quasi-quadrate), if the permutation $f|stran\ f$ has the second iterative root in the group *G(stran f)*.

Theorem 15: If $f \in T(X)$ has the second iterative root in *T(X)*, then *f* is the quasi-quadratic element in *T(X)*.

Example 1: Let $X = \{1,2,\ldots,9\}$, the mapping $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 3 & 4 & 3 & 8 & 9 & 8 & 9 & 7 \end{pmatrix}$, the vertex graph is shown in Fig. 9. $Stran\ f = f^2(X) = \{3, 4, 7, 8, 9\}$, $cont\ f = 2$, $f|stran\ f = (3\ 4).(7\ 8\ 9)$. As $f|stran\ f$ does not have the second iterative root (only one 2-cycle), *f* is not the quasi-quadrate and therefore the second iterative root does not exist.

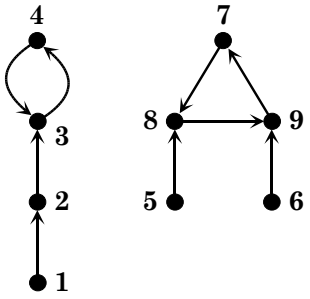


Fig 9: Vertex graph of mapping *f* from Example 1.

The previous Theorem 15 cannot be reversed. Nevertheless, the reversed theorem applies in the special case.

Theorem 16: Let $f \in T(X)$, *cont* $f = 1$. Then f has the second iterative root in $T(X)$ if and only if f is the quasi-quadratic element in $T(X)$.

Theorem 15 is only the necessary condition for the existence of the second iterative root, Theorem 16 determines the necessary and sufficient condition with the premise *cont f =1*. The general necessary and sufficient condition is given in the article [13]. We will now give this condition and supply it with examples applying the general theory. Let us remind that $h(x)$ denoted the depth of the element x below the cycle and further that examining the existence of the second iterative root of the transformation f can be performed only for quasi-quadratic transformations (Theorem 15). For every discussed transformation f of the finite set X we can suppose that $f|stran\ f$ has the second iterative root.

Definition: Let f be the transformation of the finite set X, let $x \in X - f(X)$ be an arbitrary element. Then the element $y \in X - f(X)$ is called γ-dual to the element x, if there exists such the second iterative root γ of the set $f|stran\ f$ for which there applies one of the following four conditions:

(1) $h(x) = h(y) \quad \wedge \quad \gamma[\, f^{h(x)}(x)\,] = f^{h(y)}(y)$,
(2) $h(x) = h(y) \quad \wedge \quad \gamma[\, f^{h(y)}(y)\,] = f^{h(x)}(x)$,
(3) $h(y) = h(x) + 1 \wedge \quad \gamma[\, f^{h(x)}(x)\,] = f^{h(y)}(y)$,
(4) $h(y) = h(x) - 1 \wedge \quad \gamma[\, f^{h(y)}(y)\,] = f^{h(x)}(x)$.

It is evident that the relation γ duality is symmetric; then it is possible to consider the elements x, y as mutually γ dual regardless of the order. All pairs of the γ dual elements are the elements of the set $X - f(X)$, called basic elements. Precisely speaking, the element x is the basic element if and only if it suffices the condition $(f^{-1} \circ f)(x) \cap f(X) = \phi$.

Definition: We wil say that the transformation f of the finite set X is *amenable* if it is the quasi-quadratic element in $T(X)$ and if there exists the iterative root γ of the restriction $f|stran\ f$ with the property that to each basic element of the set X there exists the γ dual element.

We have already given the definition of the equivalence π_f corresponding to the transformation f. The set of all basic elements can be decomposed into blocks of mutually equivalent basic elements. All elements in each of these blocks have the same image, the same depth below the cycle, and therefore the same γ dual element. In the next considerations we will always choose one element from each of the blocks of equivalent basic elements. The set of the chosen elements will be denoted as $B(f)$. Let γ be the second iterative root of the transformation $f|stran\ f$. On the set $B(f)$ let us define the mapping $\Delta: B(f) \rightarrow X - f(X)$ as follows: Every $x \in B(f)$ will be assigned the element which is γ dual to it in $X - f(X)$. The mapping Δ will be called the dualizing mapping because in fact for each pair of elements $(x,y) \in \Delta$ there holds that they are γ-dual. Now, for each pair $(x,y) \in \Delta$ let us denote two "iteration routes":

$(x,y)^{(1)} = \{(x, y), (y, f(x)), (f(x), f(y)), (f(y), f^2(x)), (f^2(x), f^2(y), \ldots\}$,

$(x,y)^{(2)} = \{(y, x), (x, f(y), (f(y), f(x)), (f(x), f^2(y)), (f^2(y), f^2(x)), \ldots\}$.

As the set X is finite, the number of elements of the sets $(x,y)^{(1)}$, $(x,y)^{(2)}$ is finite. Let A be any subset of the mapping Δ. Let us now define the relation Δ_A as follows:

$\Delta_A = \bigcup_{(x,y)\in A}(x,y)^{(1)} \cup \bigcup_{(x,y)\in \Delta - A}(x,y)^{(2)}$. If there exists the subset $A \subseteq \Delta$, for which the relation Δ_A is unambiguous, then the mapping Δ is called the compatibly dualizing mapping.

Definition: Let us say that the transformation f of the finite set X is *compatibly amenable* if it is amenable and there exists the compatibly dualizing mapping Δ.

Theorem 17: Let X be a finite set, let $f \in T(X)$. Then f has the second iterative root $\Leftrightarrow$ f is compatibly amenable.

Using this Theorem, the proof of which can again be found in [13], it is possible to decide theoretically unambiguously if the given transformation of the finite set has the second iterative root. However, the given theorem is quite complicated and its practical usage is possible only for transformations of finite sets with a small number of elements. From the didactic point of view, let us now show an example of the application of Theorem 17. We will show that the obtained relation Δ_A is exactly the base of the second iterative root of the transformation f, so it is possible to use Theorem 17 not only for solving the question of the existence of the second iterative root, but also for its construction. The next Theorem is useful while searching for the compatibly dualizing mapping Δ:

Theorem 18: a) The compatibly dualizing mapping Δ is bijective (it cannot be reversed).

b) Let x_1, x_2, y_1, y_2 be different elements of $B(f)$ such that every element y_1, y_2 is γ dual to every element x_1, x_2. If the mapping Δ contains pairs (x_1, y_1), $(x_2,, y_2)$, (y_1, x_2), $(y_2,,x_1)$, then it is not compatibly dualizing.

Example 2: Let $X = \{1,2,3,...,18,19\}$, let the transformation f be determined by the matrix:

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 \\ 3 & 4 & 1 & 2 & 7 & 7 & 1 & 9 & 10 & 3 & 14 & 14 & 14 & 15 & 2 & 17 & 18 & 19 & 4 \end{pmatrix}.$$

The vertex graph of the transformation f is in Fig. 10.

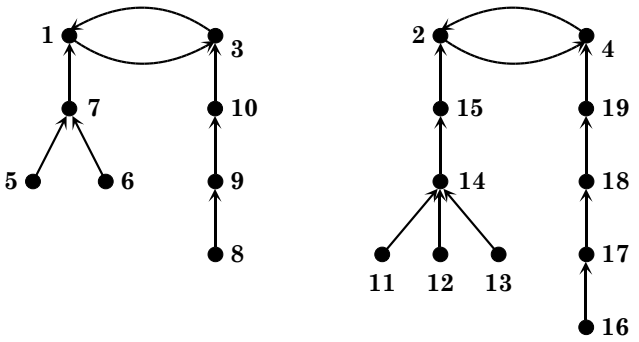


Fig. 10: Vertex graph of mapping f from Example 2.

The transformation f is the quasi-quadratic element in $T(X)$, because it contains just two cycles of the order 2 (Theorem 14). These cycles are (1, 3), (2, 4), *cont* f = 4, *stran* f = {1,2,3,4}, the iterative quasi-quadratic root γ on *stran* f is defined by the matrix $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$. The classes of equivalent basic elements are {5,6}, {8}, {11,12, 13}, {16}.

Each of the elements of the set {5, 6} is γ dual to each of the elements {11, 12, 13}, elements 8 and 16 are also γ dual. Therefore, the mapping *f* is amenable. The set *B(f)* will be chosen as {5, 8, 11, 16}. The dualizing mapping Δ is defined by the matrix $\begin{pmatrix} 5 & 8 & 11 & 16 \\ 11 & 16 & 5 & 8 \end{pmatrix}$. Now let us describe the routes for all pairs of elements of the mapping Δ. As we will soon find out, in fact it is enough to describe the routes with the exponent (1):

$(5,11)^{(1)} = (11,5)^{(2)}$ = {(5,11), (11,7), (7,14), (14,1), (1,15), (15,3), (3,2), (2,1), (1,4), (4,3)},

$(11,5)^{(1)} = (5,11)^{(2)}$ = {(11,5), (5,14), (14,7), (7,15), (15,1), (1,2), (2,3), (3,4), (4,1)},

$(8,16)^{(1)} = (16,8)^{(2)}$ = {(8,16), (16,9), (9,17), (17,10), (10,18), (18,3), (3,19), (19,1), (1,4), (4,3), (3,2), (2,1)},

$(16,8)^{(1)} = (8,16)^{(2)}$ = {(16,8), (8,17), (17,9), (9,18), (18,10), (10,19), (19,3), (3,4), (4,1), (1,2), (2,3)}.

The relations $(11,5)^{(1)}$ a $(16,8)^{(1)}$ are unambiguous, so we will denote *A* = {(11,5), (16,8)}. The relation Δ_A ={(11,5), (5,14), (14,7), (7,15), (15,1), (1,2), (2,3), (3,4), (4,1), (16,8), (8,17), (17,9), (9,18), (18,10), (10,19), (19,3)} is unambiguous, so Δ is the compatibly dualizing mapping and *f* is compatibly amenable. The desired second iterative root *g* of the transformation *f* is defined as $\gamma \cup \Delta_A$. It is again given by the matrix (and illustrated by Fig. 11).

$$g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 \\ 2 & 3 & 4 & 1 & 14 & 14 & 15 & 17 & 18 & 19 & 5 & 5 & 5 & 7 & 1 & 8 & 9 & 10 & 3 \end{pmatrix}.$$

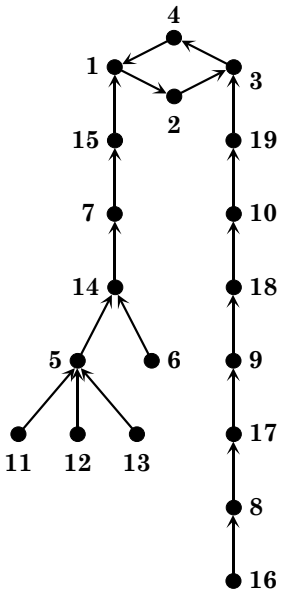


Fig. 11: Vertex graph of mapping *g*. There holds $g^2 = f$.

CONCLUSION

The article introduces the general theory of iterations of set transformations and gives basic theorems describing the questions of the existence and construction of their iterative roots; further there is stated the necessary and sufficient condition for the existence of the second iterative root of mappings defined on finite sets. The text is complemented with the possible usage of the above described theory while solving problems from the high school mathematics. From the didactic point of view, it is possible to conclude that despite a certain formal complexity (the proofs are quite complicated), these problems can be solved with talented students, and that understanding of the essence of the given theory can develop the students´ mathematical abilities and thinking processes. Among others, the ability to "deciphre" and study the formally complicated mathematical text is also extremely significant. The familiarity with vertex graphs of mappings of finite sets (including vertex graphs of functions defined on infinite sets) contributes to the better understanding of the substance of relations, mappings and functions in the classical continuous approach.

REFERENCES

[1] BERÁNEK, J., CHVALINA, J.: *On Tabor`s problem concerning a certain quasi-ordering of iterative roots of functions.* Aequ. Math. 39 (1990), pp. 1-5. ISSN 0001-9054/90/. 1990

[2] BERÁNEK, J., CHVALINA, J.: *O iteračních odmocninách kvadratické funkce*. In: Sborník prací pedagogické fakulty UJEP, řada matematických věd č. 1, Brno 1990, pp. 7-19. ISBN 80-210-0136-4

[3] BERÁNEK, J.: *O iterativních kořenech jisté polynomické funkce.* In: Sborník prací pedagogické fakulty Masarykovy univerzity, řada matematických věd č. 3, Brno 1993, pp. 5-15. ISBN 80-210-0589-0

[4] BERÁNEK, J., CHVALINA, J.: *Iterativní teorie funkcí a školská matematika.* In: Acta Fac. Paed. Univ. Tyrnaviensis, 2002, no 6.Trnavská Universita, Trnava 2002, pp. 5-10. ISBN 80-89074-51-0

[5] BERÁNEK, J.: *Hyperbolické funkce z hlediska diskrétní iterační teorie*. In: Acta mathematica 6, Sborník příspěvků z nitranské matematické konference, UKF, fakulta prírodných vied, Nitra 2003, pp. 123-135, ISBN 80-8050-666-3.

[6] BERÁNEK, J.: *Funkcionální rovnice.* Brno: Masarykova univerzita v Brně, 2004, 74 pp., ISBN 80-210-3422-X.

[7] BERÁNEK, J.: *Metrické prostory a kvadratická funkce.* In: Acta mathematica 7, Sborník příspěvků ze II. nitrianské matematické konference, UKF, fakulta prírodných vied, Nitra 2004, pp. 123-135. ISBN 80-8050-666-3.

[8] BERÁNEK, J.: *Iterativní kořeny transformací konečných množin.* In: Acta mathematica 8, první. Nitra : Fakulta prírodných vied UKF v Nitre, 2005. pp. 85-94. ISBN 80-8050-896-8

[9] CHVALINA, J.: *Funkcionální grafy, kvaziuspořádané množiny a komutativní hypergrupy*. Masarykova univerzita, Brno 1995, 205 pp., ISBN 80-210-1148-3.

[10] KUCZMA, M.: *Functional Equations in a Single Variable.* PWN, Warszawa 1968.

[11] LOJASIEWICZ, S.: *Solution générale de l`équation fonctionnelle f(f(...f(x)...)) = g(x).* Ann. Soc. Polon. Math. 24 (1951), pp. 88-91.

[12] SMÍTAL, J.: *O funkciách a funkcionálnych rovniciach*. 1st edition, Bratislava : Alfa, 1984. 143 pp.

[13] SNOWDON, M., HOWIE, J.M.: *Square roots in finite full tranformation semigroups*. Glasgow Math. J. 23 (1982), pp. 137-149.

[14] TARGONSKI, G.: *Topics in Iteration Theory*. Vandenhoeck et Ruprecht, Göttingen and Zürich 1981.

Aggregation-Disaggregation Approach for Computing the Mean First Passage Times Matrices

František Bubeník, Petr Mayer

Faculty of Civil Engineering, Czech Technical University in Prague,
Thakurova 7, 166 29 Praha 6. Czech Republic
bubenik@fsv.cvut.cz, petr.mayer@fsv.cvut.cz

Abstract: *The mean first passage times matrix (MFPTM) is one of the principal characteristics of Markov chains. Direct algorithms for its computing are known. The first one was introduced by C. D. Meyer and later M. Neumann brought some improvements. Other enhancements and a reduction in the number of operations come from P. Mayer. The first and second approaches require the inversion of a full matrix of size n or of two matrices of size n/2, respectively. The third approach inverts a sparse matrix of size n with taking advantage of an appropriate LU decomposition. A problem with efficiency occurs in particular in the case that only a small part of the MFPTM is required, because all the elements of the matrix are necessary to be determined, in principle. An iterative aggregation-disaggregation method (IAD) is successfully used for computing stationary probability distributions. This paper deals with the use of an IAD method for computing a part of the MFPTM. Conditions under which the IAD method can be used, are examined.*

Keywords: Markov chains, mean first passage times matrices, iterative aggregation-disaggregation methods, numerical methods.

Introduction

The basic motivation for the study of homogeneous Discrete Time Markov Chains (DTMC) is a quantitative risk and reliability analysis for Railways signaling systems, see [6] and [4], [5]. The probability characteristic of the transitions to these classes is the issue of the risk analysis. This is the reason why it is necessary for us to study DTMC. In this paper we confine our considerations to irreducible homogeneous finite DTMC.

We show some possibilities for computing stationary probability vectors in the case of an irreducible transition matrix and one method for computing the MFPTM.

What is new in this paper: Firstly, generally use aggregation-disaggregation algorithm for calculating columns of MFPTM (just aggregation-disaggregation approach is emphasized, column access is already described in [8]). Further, aggregate calculation of a block of MFPTM in case of a certain form of the transition matrix.

1 Basic Concepts and Characteristics of DTMC

The symbol $\mathbf{E}$ is used for the matrix of all ones and $\mathbf{e}$ for the column vector with all elements equal to 1. The dimensions of $\mathbf{E}$ and $\mathbf{e}$ will always be clear from the context. Let for any matrix $\mathbf{Y} \in \Re^{n\times n}$, $\mathbf{Y_d}$ denote the $n \times n$ diagonal matrix whose diagonal entries are the corresponding diagonal entries of $\mathbf{Y}$.

Definition 1 *Let elements of* $\mathbf{T} \in \mathfrak{R}^{n\times n}$ *be non negative and* $\mathbf{T}\,\mathbf{e} = \mathbf{e}$*, where* $\mathbf{e} = (1,\dots,1)^{\mathrm{T}} \in \mathfrak{R}^n$*. Then we say that* $\mathbf{T}$ *is a stochastic matrix.*

Definition 2 *A finite Markov chain is a stochastic process moving through a finite number of states and for which the probability of entering a certain state depends only on the last state occupied.*

Suppose that $\{X_m \,|\, m = 0,1,\dots\}$ is a finite homogeneous Markov chain on the states $S_1,\dots,S_n$. Let $\mathbf{T} \in \mathfrak{R}^{n\times n}$ be its corresponding transition matrix. More information on stochastic processes and Markov chains can be found in [1], [11]. From our point of view we are interested in times which are necessary for transitions from a state to a state.

Definition 3 *Let* $\mathbf{T}$ *be a stochastic matrix. The vector* $\pi \in \mathfrak{R}^n$ *is called the stationary probability vector (SPV) if* $\pi^{\mathrm{T}} = \pi^{\mathrm{T}}\mathbf{T}$*,* $\pi^{\mathrm{T}}\mathbf{e} = 1$.

The existence and uniqueness of SPV are studied in [1], [11].

We introduce the probabilities of changes from a state to a state in terms of matrices as follows

Definition 4 *We denote by* $f_{ij}^{(l)}$ *the probability of the first come to the state* j *after leaving the state* i *and it occurs exactly in* l *time steps. Let us denote by* f_{ij} *the total probability of the transition from the state* i *to the state* j*, i. e.* $f_{ij} = \sum_{l=1}^{\infty} f_{ij}^{(l)}$*. Let us define* $\mathbf{F} = (f_{ij})_{i,j=1}^{n}$ *and* $\mathbf{F}^{(l)} = (f_{ij}^{(l)})_{i,j=1}^{n}$.

The correctness of the definition f_{ij} follows from the independence of events with probabilities represented by $f_{ij}^{(l)}$. Situation when f_{ij} is equal to 1 occurs only in a special case. In our case we are considering only irreducible chains, therefore, f_{ij} is equal to 1 always occurs.

Let us introduce a model example illustrating the theme. Consider a chain as in Figure 1.

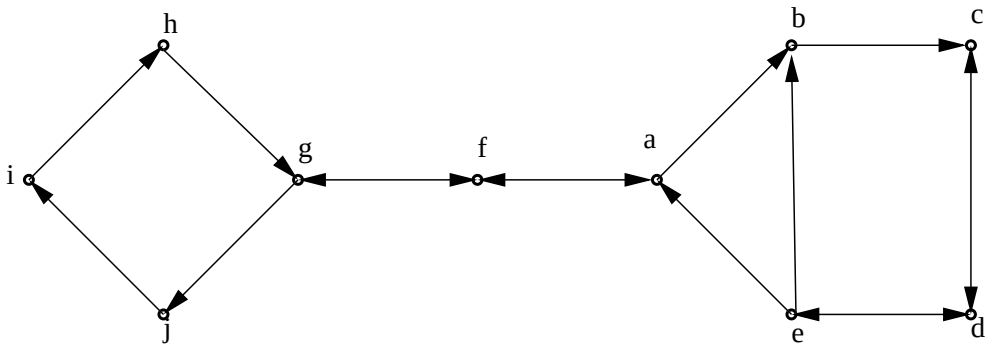


Fig. 1: The Markov chain.

with the transition matrix $\mathbf{T}$

$$\mathbf{T} = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \\ j \end{array} \begin{array}{c} \begin{array}{cccccccccc} a & b & c & d & e & f & g & h & i & j \end{array} \\ \begin{pmatrix} \cdot & 0.2 & \cdot & \cdot & \cdot & 0.8 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.3 & \cdot & 0.7 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0.5 & \cdot & 0.5 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0.4 & \cdot & \cdot & 0.6 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0.1 & \cdot & \cdot & \cdot & \cdot & \cdot & 0.9 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0.8 & \cdot & \cdot & \cdot & 0.2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \end{pmatrix} \end{array} \tag{1}$$

Then the corresponding SPV is

$$\pi = \big(\; 0.039620, \quad 0.007924, \quad 0.016640, \quad 0.028526, \quad 0.019810,$$
$$0.316957, \quad 0.356377, \quad 0.071315, \quad 0.071315, \quad 0.071315 \; \big) \tag{2}$$

According to a well-known theorem of Kolmogorov, $(\mathbf{T}^p)_{ij}$ is the probability that the transition from the $i-$th state to the $j-$th state occurs just (but not necessarily for the first time) in the $p-$th step. For $l \leq p$, consider the probability that the transition from the state i to the state j occurs for the first time just in the $l-$th step and then in $l-p$ steps the transition follows back to the state j.

We add it for l from 1 to p and then we can write (see also [11])

$$(\mathbf{T}^p)_{ij} = \sum_{l=1}^{p} f_{ij}^{(l)} \left(\mathbf{T}^{(p-l)}\right)_{jj}$$

which implies

$$f_{ij}^{(p)} = (\mathbf{T}^p)_{ij} - \sum_{l=1}^{p-1} f_{ij}^{(l)} \left(\mathbf{T}^{(p-l)}\right)_{jj}.$$

It is in the matrix form

$$\mathbf{T}^p = \sum_{l=1}^{p} \mathbf{F}^{(l)} \left(\mathbf{T}^{(p-l)}\right)_{\mathbf{d}}, \tag{3}$$

$$\mathbf{F}^{(p)} = \mathbf{T}^p - \sum_{l=1}^{p-1} \mathbf{F}^{(l)} \left(\mathbf{T}^{(p-l)}\right)_{\mathbf{d}}. \tag{4}$$

The total probability of the transition from the state i to the state j can be considered as the sum of the probability of the transition from the state i to the state j in one step and the sum of probabilities of the transition from the state i to the state k, different from j, in one step and from the state k to the state j. Then we can write

$$f_{ij} = t_{ij} + \sum_{k \neq j} t_{ik} \, f_{kj}, \tag{5}$$

since $\mathbf{T}$ is the transition matrix of a homogeneous Markov chain and the formula (5) represents one step of the chain applied from the state i to the state j.

Rewritten in the matrix form, we obtain

$$\mathbf{F} = \mathbf{T} + \mathbf{TF} - \mathbf{TF_d} = \mathbf{T}(\mathbf{F} - \mathbf{F_d}) + \mathbf{T}. \tag{6}$$

If the transition matrix $\mathbf{T}$ is irreducible then $\mathbf{F} = \mathbf{E}$ and it expresses the sure event. In other words, any transition is possible.

Definition 5 *We denote by m_{ij} the mean first passage time of the transition from the state i to the state j, i. e. $m_{ij} = \sum_{l=1}^{\infty} l f_{ij}^{(l)}$. Let us define $\mathbf{M} = (m_{ij})_{i,j=1}^{n}$.*

When we read Definition 5, from another point of view, we can see that each element m_{ij}, which is called the first statistical moment of the transition, is equal to the weighted mean of the lengths with their relative frequencies, which are their probabilities, as their weights.

In a similar way as for $\mathbf{F}$ in (6), we get a formula for computing the MFPTM in the matrix form

$$\mathbf{M} = \mathbf{T} + \mathbf{T}(\mathbf{F} - \mathbf{F_d}) + \mathbf{T}(\mathbf{M} - \mathbf{M_d})$$

and if we apply the identity $\mathbf{T} + \mathbf{T}(\mathbf{F} - \mathbf{F_d}) = \mathbf{F}$ from (6), we get

$$\mathbf{M} = \mathbf{T}(\mathbf{M} - \mathbf{M_d}) + \mathbf{F}. \tag{7}$$

To our knowledge, if $\mathbf{T}$ is irreducible, i. e. $\mathbf{F} = \mathbf{E}$ is the matrix of all $1s$, the previous equality is well known. For a stationary vector π

$$\pi\,\mathbf{M_d} = \pi\,\mathbf{F}.$$

In case of $\mathbf{T}$ irreducible, the last expression is just the well-known renewal theorem. In such case, there is $\pi_i m_{ii} = 1$, i. e.

$$m_{ii} = \frac{1}{\pi_i}.$$

According to Meyer [9], the mean first passage matrix $\mathbf{M}$ (note that if there is necessary to emphasize that $\mathbf{M}$ corresponds to the transition matrix $\mathbf{T}$ we use the notation $\mathbf{M_T}$) is given by

$$\mathbf{M} = [\mathbf{I} - \mathbf{Q}^{\#} + \mathbf{E}\,\mathbf{Q}_{\mathbf{d}}^{\#}]\mathbf{\Pi}^{-1}, \tag{8}$$

where $\mathbf{Q} = \mathbf{I} - \mathbf{T}$, where $\mathbf{Q}^{\#}$ is the group (generalized) inverse of $\mathbf{Q}$, i. e.

$$\mathbf{Q}^{\#} = (\mathbf{I} - \mathbf{T})^{\#} = (\mathbf{I} - \mathbf{T} + \mathbf{e}\,\pi^{\mathrm{T}})^{-\mathbf{1}} - \mathbf{e}\,\pi^{\mathrm{T}}$$

and where $\mathbf{\Pi}$ is the diagonal matrix whose diagonal entries are the corresponding entries of π.

The mean first passage times matrix for our model chain is

$$\mathbf{M} = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \\ j \end{array} \begin{array}{c} \begin{array}{ccccc} a & b & c & d & e \end{array} \\ \left(\begin{array}{ccccc} 25.240000 & 117.000000 & 199.333333 & 165.277777 & 119.200000 \\ 9.200000 & 126.200000 & 82.333333 & 48.277777 & 2.200000 \\ 11.000000 & 128.000000 & 60.095238 & 1.000000 & 4.000000 \\ 10.000000 & 127.000000 & 59.095238 & 35.055555 & 3.000000 \\ 7.000000 & 124.000000 & 116.190476 & 67.111111 & 50.480000 \\ \hline 28.000000 & 145.000000 & 227.333333 & 193.277777 & 147.200000 \\ 30.000000 & 147.000000 & 229.333333 & 195.277777 & 149.200000 \\ 31.000000 & 148.000000 & 230.333333 & 196.277777 & 150.200000 \\ 32.000000 & 149.000000 & 231.333333 & 197.277777 & 151.200000 \\ 33.000000 & 150.000000 & 232.333333 & 198.277777 & 152.200000 \end{array}\right) \end{array}$$

$$\left(\begin{array}{|ccccc|}
f & g & h & i & j \\
\hline
3.550000 & 5.055555 & 18.077777 & 17.077777 & 16.077777 \\
12.750000 & 14.255555 & 27.277777 & 26.277777 & 25.277777 \\
14.550000 & 16.055555 & 29.077777 & 28.077777 & 27.077777 \\
13.550000 & 15.055555 & 28.077777 & 27.077777 & 26.077777 \\
10.550000 & 12.055555 & 25.077777 & 24.077777 & 23.077777 \\
\hline
3.155000 & 1.505555 & 14.527777 & 13.527777 & 12.527777 \\
2.000000 & 2.804444 & 13.022222 & 12.022222 & 11.022222 \\
3.000000 & 1.000000 & 14.022222 & 13.022222 & 12.022222 \\
4.000000 & 2.000000 & 1.000000 & 14.022222 & 13.022222 \\
5.000000 & 3.000000 & 2.000000 & 1.000000 & 14.022222
\end{array}\right) \tag{9}$$

2 IAD Algorithm

We introduce an aggregation mapping

$$g : \{1, \ldots, N\} \to \{1, \ldots, n\}, \quad n \ll N,$$

where n is the size of the coarse space.

The indices which are mapped to the same values of g define one aggregation group. The optimal choice of mapping g is difficult and often depends on further information about the solved problem. Distinctions between two choices of g for the same transition matrix can be substantial.

Consider the aggregation mapping

$$\begin{aligned} g : \{1,2,3,4,5\} &\to 1, \quad g : 6 \to 2, \quad g : 7 \to 3, \\ g : 8 &\to 4, \quad g : 9 \to 5, \quad g : 10 \to 6. \end{aligned} \tag{10}$$

By means of aggregation mappings we define the restriction and prolongation matrices.

The restriction matrix $\mathbf{R} \in \mathfrak{R}^{N \times n}$ is defined by nonzero elements

$$r_{g(i),i} = 1,$$

i. e. $(\mathbf{R}\,\mathbf{x})_j = \sum_{i=1, g(i)=j}^{N} x_i$.

The restriction matrix to the model chain is

$$\mathbf{R} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{11}$$

The prolongation matrix $\mathbf{S}(\mathbf{x})$ is parameterized by a vector $\mathbf{x} \in \mathfrak{R}^{\mathbf{N}}$; the nonzero elements of the matrix are

$$(\mathbf{S}(\mathbf{x}))_{i,g(i)} = \frac{x_i}{(\mathbf{R}\,\mathbf{x})_{g(i)}},$$

it means that $(\mathbf{S}(\mathbf{x})\,\mathbf{z})_i = z_{g(i)}\, x_i / (\mathbf{R}\,\mathbf{x})_{g(i)}$.

The prolongation matrix for the model chain is

$$\mathbf{S}(\mathbf{x}) = \begin{pmatrix} 0.352113 & 0 & 0 & 0 & 0 & 0 \\ 0.070043 & 0 & 0 & 0 & 0 & 0 \\ 0.147887 & 0 & 0 & 0 & 0 & 0 \\ 0.253521 & 0 & 0 & 0 & 0 & 0 \\ 0.176056 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{12}$$

Let us denote by $\mathbf{A}(\mathbf{x}) = \mathbf{R}\,\mathbf{T}^{\mathrm{T}}\,\mathbf{S}(\mathbf{x})$ the aggregated matrix defined by the vector $\mathbf{x}$ and by the aggregation mapping g. Some properties of the matrix $\mathbf{A}(\mathbf{x})$ are introduced in the following lemma.

Lemma 1 *Let $\mathbf{T}$ be a stochastic matrix, let g be an aggregation mapping and $\mathbf{x} \in \mathfrak{R}^{\mathbf{N}}$ such that $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{R}\,\mathbf{x} > \mathbf{0}$. Then the aggregated matrix $\mathbf{A}(\mathbf{x})$ is a column stochastic matrix. If the matrix $\mathbf{T}$ is irreducible and the vector $\mathbf{x}$ is strictly positive, then $\mathbf{A}(\mathbf{x})$ is irreducible.*

With the previous knowledge we can define the following IAD algorithm for an irreducible stochastic matrix $\mathbf{T}$ and for a positive initial approximation $\mathbf{x_{init}}$. Suppose that matrices $\mathbf{W_1}$ and $\mathbf{W_2}$ form the regular splitting of the matrix $\mathbf{I} - \mathbf{T}^{\mathrm{T}}$. It means that $\mathbf{I} - \mathbf{T}^{\mathrm{T}} = \mathbf{W_1} - \mathbf{W_2}$, where $\mathbf{W_1}$ is a $M-$matrix and where $\mathbf{W_2}$ is a nonnegative matrix.

Algorithm IAD (input: $\mathbf{T}$, $\mathbf{W_1}$, $\mathbf{W_2}$, $\mathbf{x_{init}}$, ε, g, s; output: $\mathbf{x}$)
1. $k := 1$, $\mathbf{x_1} := \mathbf{x_{init}}$
2. while $||\mathbf{T}^{\mathrm{T}}\,\mathbf{x}_k - \mathbf{x}_k|| > \varepsilon$ do
3. $\quad \widetilde{\mathbf{x}} := (\mathbf{W_1^{-1}}\,\mathbf{W_2})^s\,\mathbf{x}_k$
4. $\quad \mathbf{A}(\widetilde{\mathbf{x}}) := \mathbf{R}\,\mathbf{T}^{\mathrm{T}}\mathbf{S}(\widetilde{\mathbf{x}})$
5. $\quad$ solve $\mathbf{A}(\widetilde{\mathbf{x}})\,\mathbf{z} = \mathbf{z}$ and $\mathbf{e}^{\mathrm{T}}\,\mathbf{z} = 1$
6. $\quad k := k + 1$
7. $\quad \mathbf{x}_k = \mathbf{S}(\widetilde{\mathbf{x}})\,\mathbf{z}$
8. end while

The convergence theory for IAD can be found in [7].

3 Alternative Computation of M

3.1 Algorithm for Computing M by Parts

This algorithm is included here for the sake of completeness. In full, this is introduced in [3].

Consider an irreducible stochastic transition matrix $\mathbf{T}$ of order n. We assume, without loss of generality, that $\mathbf{T}$ has the partitioned form

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix}, \tag{13}$$

where $\mathbf{T}_{11}$ and $\mathbf{T}_{22}$ are square matrices such that the sum of their orders gives n and the other blocks are submatrices of corresponding dimensions. A vector $\mathbf{x} \in \Re^n$ is partitioned into blocks $\mathbf{x}_1$ and $\mathbf{x}_2$ conformably to partitioning of $\mathbf{T}$.

The Perron complement of $\mathbf{T}_{11}$ in $\mathbf{T}$ is given by

$$\mathcal{P}(\mathbf{T}/\mathbf{T}_{11}) = \mathbf{T}_{11} + \mathbf{T}_{12}(\rho(\mathbf{T})\mathbf{I} - \mathbf{T}_{22})^{-1}\,\mathbf{T}_{21}, \tag{14}$$

where $\rho(\cdot)$ denotes the spectral radius of a matrix. Note that from the well-known Perron-Frobenius Theorem we know that $\rho(\mathbf{T}) = 1$. (Since $\mathbf{T}$ is irreducible, $\rho(\mathbf{T}) > \rho(\mathbf{T}_{22})$, so that the expression on the right hand side of (14) is well defined.) For more details see [3]. Note that as $\mathbf{T}$ is supposed to be irreducible, then all complements are stochastic and irreducible matrices.

We denote the Perron complements $\mathcal{P}(\mathbf{T}/\mathbf{T}_{11})$ and $\mathcal{P}(\mathbf{T}/\mathbf{T}_{22})$ by $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively (note that these matrices are of the same orders as $\mathbf{T}_{11}$ and $\mathbf{T}_{22}$, respectively).

Recall from Meyer [10] that if ξ_1 is the stationary probability vector for $\mathcal{P}_1$, then

$$\gamma_1\pi_1 = \xi_1, \tag{15}$$

where $1/\gamma_1 = \mathbf{e}^{\mathrm{T}}\,\pi_1$. Analogously for ξ_2.

For the sake of simplicity in describing the method, assume that the transition matrix $\mathbf{T}$ is partitioned as in (13) and assume that the mean first passage matrix $\mathbf{M}$ is partitioned conformably as

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix}. \tag{16}$$

The following Theorem is based on Theorems 2.2 and 2.3 from [3].

Theorem 1 *Let $\mathbf{T} \in \Re^{n\times n}$ be a non-negative stochastic and irreducible matrix and let $\mathbf{M}$ be its corresponding mean first passage matrix. Partition $\mathbf{T}$ as in (13). Then*

$$\mathbf{M}_{11} = \gamma_1(\mathbf{M}_{\mathcal{P}_1}) + \mathbf{V}_1,\ \mathbf{M}_{22} = \gamma_2(\mathbf{M}_{\mathcal{P}_2}) + \mathbf{V}_2,$$

where γ_1 and γ_2 are determined via (15) and where $\mathbf{V}_1$ and $\mathbf{V}_2$ are the skew symmetric matrices (of rank at most 2) given by

$$\mathbf{V}_1 = (\mathbf{I} - \mathcal{P}_1)^{\#}\mathbf{T}_{12}(\mathbf{I} - \mathbf{T}_{22})^{-1}\mathbf{E} - ((\mathbf{I} - \mathcal{P}_1)^{\#}\mathbf{T}_{12}(\mathbf{I} - \mathbf{T}_{22})^{-1}\mathbf{E})^{\mathrm{T}}, \tag{17}$$

$$\mathbf{V}_2 = (\mathbf{I} - \mathcal{P}_2)^{\#}\mathbf{T}_{21}(\mathbf{I} - \mathbf{T}_{11})^{-1}\mathbf{E} - ((\mathbf{I} - \mathcal{P}_2)^{\#}\mathbf{T}_{21}(\mathbf{I} - \mathbf{T}_{11})^{-1}\mathbf{E})^{\mathrm{T}}. \tag{18}$$

Further

$$\mathbf{M}_{21} = (\mathbf{I} - \mathbf{T}_{22})^{-1}\,\mathbf{T}_{12}\,[\mathbf{M}_{22} - (\mathbf{M}_{22})_{\mathbf{d}}] + (\mathbf{I} - \mathbf{T}_{22})^{-1}\,\mathbf{E} \tag{19}$$

and

$$\mathbf{M}_{12} = (\mathbf{I} - \mathbf{T}_{11})^{-1}\mathbf{T}_{21}[\mathbf{M}_{11} - (\mathbf{M}_{11})_{\mathbf{d}}] + (\mathbf{I} - \mathbf{T}_{11})^{-1}\,\mathbf{E}. \tag{20}$$

The proof can be found in [3].

The algorithm consists in two steps:

Step (i): Compute the diagonal blocks of $\mathbf{M}$. Taking the diagonal block $\mathbf{M}_{11}$ as a representative, we see (from Theorem 1) that in order to find $\mathbf{M}_{11}$, we must have γ_1, V_1 and $\mathbf{M}_{\mathcal{P}_1}$. We

first need $\mathcal{P}_1$, from which we can find both ξ_1 and $(\mathbf{I}-\mathcal{P}_1)^{\#}$. Having found $(\mathbf{I}-\mathcal{P}_1)^{\#}$, we then use it and ξ_1 to compute both V_1 (from (17)) and $\mathbf{M}_{\mathcal{P}_1}$ (from (8)).

An analogous set of computations can be performed independently in order to obtain γ_2, V_2 and $\mathbf{M}_{\mathcal{P}_2}$. With ξ_1 and ξ_2 in hand, we use the fact that $1/\gamma_1$ and $1/\gamma_2$ form, respectively, the entries of the normalized left Perron vector of the 2×2 coupling matrix

$$\mathbf{C} = \begin{pmatrix} \xi_1^{\mathrm{T}}\,\mathbf{T}_{11}\,\mathbf{e} & \xi_1^{\mathrm{T}}\,\mathbf{T}_{12}\,\mathbf{e} \\ \\ \xi_2^{\mathrm{T}}\,\mathbf{T}_{21}\,\mathbf{e} & \xi_2^{\mathrm{T}}\,\mathbf{T}_{22}\,\mathbf{e} \end{pmatrix}. \tag{21}$$

Note that $\mathbf{C} = [\mathbf{A}(\pi)]^{\mathrm{T}}$, when the aggregation mapping is

$$g : \begin{cases} \{1,\ldots,n_1\} & \longrightarrow \quad 1, \\ \{n_1+1,\ldots,n\} & \longrightarrow \quad 2, \end{cases}$$

where n_1 denotes the size of $\mathbf{M}_{11}$.

Having thus found γ_1 and γ_2, we then compute $\mathbf{M}_{11}$ and $\mathbf{M}_{22}$, according to Theorem 1.

Step (ii): Compute the off-diagonal blocks of $\mathbf{M}$. From Theorem 1, this can be accomplished once we have found both $\mathbf{M}_{11}$ and $\mathbf{M}_{22}$, see (19), (20).

3.2 Computing $\mathbf{M}$ by Columns

Now, we suggest a different approach; to compute each column of $\mathbf{M}$ separately and then to apply the Sherman-Morrison-Woodbury formula. It is another variant of computing $\mathbf{M}$, based on the formula (32), also (23) and (24), which is proved in [8] and which consists in computation by columns. An effective implementation is described in [8].

Even in the worst case, this algorithm requires just n^3 operations and further improvements can be achieved if a sparse structure is available. Moreover, this approach gives an advantage if some elements of $\mathbf{M}$ only are needed, then a column or some columns can only be computed.

Denote by $\mathbf{M}_i$ the $i-$th column of the matrix $\mathbf{M}$. Rewriting (7) for the $i-$th column, we obtain

$$\mathbf{M}_i = \mathbf{e} + \mathbf{T}\,\mathbf{M}_i - (\mathbf{T}\,\mathbf{M}_{\mathbf{d}})_i, \tag{22}$$

where $(\mathbf{T}\,\mathbf{M}_{\mathbf{d}})_i \in \Re^n$ is the $i-$th column of the matrix $\mathbf{T}\,\mathbf{M}_{\mathbf{d}}$. We can write (22) as

$$\mathbf{M}_i = \mathbf{e} + \mathbf{T}\,\mathbf{M}_i - m_{ii}\,\mathbf{T}\,\mathbf{e}_i.$$

Put $\mathbf{T}_{*i} = \mathbf{T}\,\mathbf{e}_i\,\mathbf{e}_i^{\mathrm{T}}$. Then $\mathbf{T}_{*i}$ is the matrix of the same size as $\mathbf{T}$, but it has only one (the $i-$th) nonzero column and then

$$\mathbf{M}_i = \mathbf{e} + \mathbf{T}\,\mathbf{M}_i - \mathbf{T}_{*i}\,\mathbf{M}_i.$$

Thus, we can write

$$(\mathbf{I} - \mathbf{T} + \mathbf{T}_{*i})\,\mathbf{M}_i = \mathbf{e} \tag{23}$$

and if we denote $\mathbf{A}_i = \mathbf{I} - \mathbf{T} + \mathbf{T}_{*i}$, we get

$$\mathbf{A}_i\,\mathbf{M}_i = \mathbf{e}. \tag{24}$$

3.3 Aggregated Computation of a Block of M

Definition 6 *If the graph of a transition matrix* $\mathbf{T}$ *is dichotomized by omitting a vertex into two connected subgraphs, then such a vertex is called a cut point.*

The vertex f in Fig.1 is an example of a cut point. If a matrix $\mathbf{T}$ contains a cut point in the sense of Definition 6, then the individual blocks of the matrix $\mathbf{M}$ can be computed by aggregations.

In Fig.1, let us consider the vertices $a,\ b, \ldots, j$ enumerated subsequently by numbers $1,\ 2, \ldots, 10$. Then the aggregation (10) assures that the parts of the matrices $\mathbf{M}_{[\mathbf{A}(\pi)]^{\mathrm{T}}}$ in (31) and $\mathbf{M}_{22}$ from (9) corresponding to vertices f, g, h, i, j (nonaggregated parts) are equal.

This fact can be generalized into the following Theorem, the proof of which comes from the text.

Theorem 2 *Let* $\mathbf{T}$ *contain a cut point. Consider the aggregation* g *such that the vertices of one subgraph are joined with the cut point and the vertices of the other subgraph are not aggregated. Then the part of the matrix* $\mathbf{M}$ *corresponding to the nonaggregated indices coincides with the appropriate part of the original matrix* $\mathbf{M}$.

Furthermore, we formulate and prove a more general situation in which a cut point is a particular case.

Suppose that matrix $\mathbf{T} \in \Re^{n\times n}$ is partitioned as in (13), where $\mathbf{T}_{11} \in \Re^{n_1\times n_1}$, $\mathbf{T}_{12} \in \Re^{n_1\times n_2}$, $\mathbf{T}_{21} \in \Re^{n_2\times n_1}$, $\mathbf{T}_{22} \in \Re^{n_2\times n_2}$ and where $n = n_1 + n_2$. Further $\mathbf{T} \geqq 0$ and $\mathbf{T}\ \mathbf{e} = \mathbf{e}$, where $\mathbf{e} = (1, \ldots, 1)^{\mathrm{T}}$ and $\mathbf{T}$ is a stochastic matrix (see Definition 1).

We suppose that the corresponding stationary probability vector π (see Definition 3) is blocked conformably, i. e.

$$\pi = \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix}.$$

It is easy to see (from Definition 1 and Definition 3) the following important properties and relations:

Lemma 2 *It is true that*

$$\begin{aligned} \begin{pmatrix} \pi_1^{\mathrm{T}} & \pi_2^{\mathrm{T}} \end{pmatrix} &= \begin{pmatrix} \pi_1^{\mathrm{T}} & \pi_2^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix} \\ \pi_1^{\mathrm{T}} &= \pi_1^{\mathrm{T}}\mathbf{T}_{11} + \pi_2^{\mathrm{T}}\mathbf{T}_{21} \\ \pi_2^{\mathrm{T}} &= \pi_1^{\mathrm{T}}\mathbf{T}_{12} + \pi_2^{\mathrm{T}}\mathbf{T}_{22} \\ \mathbf{e}_1 &= \mathbf{T}_{11}\,\mathbf{e}_1 + \mathbf{T}_{12}\,\mathbf{e}_2 \\ \mathbf{e}_2 &= \mathbf{T}_{21}\,\mathbf{e}_1 + \mathbf{T}_{22}\,\mathbf{e}_2 \end{aligned} \tag{25}$$

Definition 7 *Matrix* $\mathbf{T}$, *partitioned as in* (13), *has the so called BM property, if the rank of block* $\mathbf{T}_{21}$ *is equal to 1, i. e.*

$$\mathbf{T}_{21} = \mathbf{u}_2\mathbf{v}_1^{\mathrm{T}}.$$

For the purposes of calculating block $\mathbf{M}_{22}$, we use matrices $\mathbf{R}$ and $\mathbf{S}$ in the following forms

$$\mathbf{R} = \begin{pmatrix} \mathbf{e}^{\mathrm{T}} & \mathbf{0} \\ \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \tag{26}$$

$$\mathbf{S} = \begin{pmatrix} \mathbf{s} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \tag{27}$$

where $\mathbf{s} \in \Re^{n_1}$, $s_i = \pi_i / (\mathbf{R}\pi)_1$ for $i = 1, \ldots, n_1$, which corresponds to the aggregation described in (10), see e. g. (11), (12).

Denote the corresponding aggregated matrix as

$$\mathbf{T}_a = \mathbf{S}^{\mathrm{T}} \mathbf{T}\, \mathbf{R}^{\mathrm{T}}, \tag{28}$$

which is in more details

$$\mathbf{T}_a = \begin{pmatrix} \frac{\pi_1^{\mathrm{T}} \mathbf{T}_{11}\, \mathbf{e}}{\pi_1^{\mathrm{T}} \mathbf{e}} & \frac{\pi_1^{\mathrm{T}} \mathbf{T}_{12}\, \mathbf{e}}{\pi_1^{\mathrm{T}} \mathbf{e}} \\ \mathbf{T}_{21}\, \mathbf{e} & \mathbf{T}_{22} \end{pmatrix}.$$

Notice that $\mathbf{T}_a$ is transposed to $\mathbf{A}(\mathbf{x})$ (see p. 40) and $\mathbf{T}_a$ is a stochastic irreducible matrix (see e. g. [7]).

It is easy to see that the left eigenvector corresponding to eigenvalue 1 of the aggregated matrix $\mathbf{T}_a$ is

$$\pi_a = \begin{pmatrix} \mathbf{e}^{\mathrm{T}} \pi_1 \\ \pi_2 \end{pmatrix}$$

(and this vector is the SPV of the matrix $\mathbf{T}_a$).

Similarly as the Perron complement of $\mathbf{T}_{11}$ in (14) we have the Perron complement of $\mathbf{T}_{22}$ in the form

$$\mathcal{P}_2 = \mathbf{T}_{22} + \mathbf{T}_{21} \left(\mathbf{I} - \mathbf{T}_{22}\right)^{-1} \mathbf{T}_{12}. \tag{29}$$

Lemma 3 *Let* $\mathbf{T}$ *have the BM property then the Perron complements* $\mathcal{P}_2$ *and* $\mathcal{P}_{a_2}$ *coincide.*

Proof: We know that

$$\mathcal{P}_2 = \mathbf{T}_{22} + \mathbf{T}_{21} \left(\mathbf{I} - \mathbf{T}_{11}\right)^{-1} \mathbf{T}_{12} = \mathbf{T}_{22} + \mathbf{u}_2 \mathbf{v}_1^{\mathrm{T}} \left(\mathbf{I} - \mathbf{T}_{11}\right)^{-1} \mathbf{T}_{12}.$$

Put

$$\mathbf{x}^{\mathrm{T}} = \mathbf{v}_1^{\mathrm{T}} (\mathbf{I} - \mathbf{T}_{11})^{-1} \mathbf{T}_{12}.$$

Then

$$\mathcal{P}_2 = \mathbf{T}_{22} + \mathbf{u}_2 \mathbf{x}^{\mathrm{T}}.$$

Similarly

$$\mathcal{P}_{a2} = \mathbf{T}_{22} + \mathbf{T}_{21}\mathbf{e} \frac{1}{\left(1 - \frac{\pi_1^{\mathrm{T}} \mathbf{T}_{11} \mathbf{e}}{\pi_1^{\mathrm{T}} \mathbf{e}}\right)} \frac{\pi_1^{\mathrm{T}} \mathbf{T}_{12}}{\pi_1^{\mathrm{T}} \mathbf{e}} = \mathbf{T}_{22} + \mathbf{u}_2 \mathbf{v}_1^{\mathrm{T}} \mathbf{e} \frac{1}{\left(1 - \frac{\pi_1^{\mathrm{T}} \mathbf{T}_{11} \mathbf{e}}{\pi_1^{\mathrm{T}} \mathbf{e}}\right)} \frac{\pi_1^{\mathrm{T}} \mathbf{T}_{12}}{\pi_1^{\mathrm{T}} \mathbf{e}}.$$

Let us denote

$$\mathbf{y}^{\mathrm{T}} = v_1^{\mathrm{T}} \mathbf{e} \frac{1}{\left(1 - \frac{\pi_1^{\mathrm{T}} \mathbf{T}_{11} \mathbf{e}}{\pi_1^{\mathrm{T}} \mathbf{e}}\right)} \frac{\pi_1^{\mathrm{T}} \mathbf{T}_{12}}{\pi_1^{\mathrm{T}} \mathbf{e}}$$

and then we have

$$\mathcal{P}_{a2} = \mathbf{T}_{22} + \mathbf{u}_2 \mathbf{y}^{\mathrm{T}}.$$

It follows from the theory of the Perron complements that the complements are stochastic matrices and the corresponding eigenvector is up to a normative constant the vector π_2. Thus

$$\pi_2^{\mathrm{T}} = \pi_2^{\mathrm{T}}\mathcal{P}_2 = \pi_2^{\mathrm{T}}\mathcal{P}_{a2}$$

and

$$\begin{aligned}
\pi_2^{\mathrm{T}}\mathbf{T}_{22} + \pi_2^{\mathrm{T}}\mathbf{u}_2\mathbf{x}^{\mathrm{T}} &= \pi_2^{\mathrm{T}}\mathbf{T}_{22} + \pi_2^{\mathrm{T}}\mathbf{u}_2\mathbf{y}^{\mathrm{T}} \\
\pi_2^{\mathrm{T}}\mathbf{u}_2\mathbf{x}^{\mathrm{T}} &= \pi_2^{\mathrm{T}}\mathbf{u}_2\mathbf{y}^{\mathrm{T}} \\
0 &= \pi_2^{\mathrm{T}}\mathbf{u}_2\mathbf{x}^{\mathrm{T}} - \pi_2^{\mathrm{T}}\mathbf{u}_2\mathbf{y}^{\mathrm{T}} \\
0 &= \pi_2^{\mathrm{T}}\mathbf{u}_2\left(\mathbf{x}^{\mathrm{T}} - \mathbf{y}^{\mathrm{T}}\right)
\end{aligned}$$

and since $\pi_2^{\mathrm{T}}\mathbf{u}_2 > 0$ it is $\mathbf{x} = \mathbf{y}$. Then also $\mathcal{P}_2 = \mathcal{P}_{a_2}$.

The principal result is formulated in the following theorem:

Theorem 3 *Let* $\mathbf{T}$ *have the BM property (see Definition 7) then the blocks* $\mathbf{M}_{22}$ *and* $\mathbf{M}_{a22}$ *coincide.*

Proof: Recall that (see Theorem 1)

$$\begin{aligned}
\mathbf{M}_{22} &= \gamma_2\mathbf{M}_{\mathcal{P}_2} + \mathbf{V}_2 \\
\mathbf{V}_2 &= \mathbf{W}_2 - \mathbf{W}_2^{\mathrm{T}} \\
\mathbf{W}_2 &= (\mathbf{I} - \mathcal{P}_2)^{\#}\,\mathbf{T}_{21}\,(\mathbf{I} - \mathbf{T}_{11})^{-1}\,\mathbf{e}\,\mathbf{e}^{\mathrm{T}} \\
\mathbf{W}_2 &= (\mathbf{I} - \mathcal{P}_2)^{\#}\,\mathbf{u}_2 v_1^{\mathrm{T}}\,(\mathbf{I} - \mathbf{T}_{11})^{-1}\,\mathbf{e}\,\mathbf{e}^{\mathrm{T}}
\end{aligned}$$

and then similarly

$$\begin{aligned}
\mathbf{M}_{a22} &= \gamma_{a2}\mathbf{M}_{\mathcal{P}_{a2}} + \mathbf{V}_{a2} \\
\mathbf{V}_{a2} &= \mathbf{W}_{a2} - \mathbf{W}_{a2}^{\mathrm{T}} \\
\mathbf{W}_{a2} &= (\mathbf{I} - \mathcal{P}_{a2})^{\#}\,\mathbf{T}_{a21}\,(\mathbf{I} - \mathbf{T}_{a11})^{-1}\,\mathbf{e}\,\mathbf{e}^{\mathrm{T}} \\
\mathbf{W}_{a2} &= (\mathbf{I} - \mathcal{P}_{a2})^{\#}\,\mathbf{u}_2\mathbf{v}_1^{\mathrm{T}}\mathbf{e}\left(1 - \frac{\pi_1^{\mathrm{T}}\mathbf{T}_{11}\mathbf{e}}{\pi_1^{\mathrm{T}}\mathbf{e}}\right)^{-1}\mathbf{e}^{\mathrm{T}}.
\end{aligned}$$

We know that $\mathcal{P}_2 = \mathcal{P}_{a2}$ (from Lemma 3) and $\gamma_2 = \gamma_{a2}$ and then it is also $\mathbf{M}_{\mathcal{P}_2} = \mathbf{M}_{\mathcal{P}_{a2}}$. We denote that $\mathbf{r} = (\mathbf{I} - \mathcal{P}_{a2})^{\#}\,\mathbf{u}_2$, $\alpha = \mathbf{v}_1^{\mathrm{T}}\,(\mathbf{I} - \mathbf{T}_{11})^{-1}\,\mathbf{e}$ and $\alpha_a = \mathbf{v}_1^{\mathrm{T}}\mathbf{e}\left(1 - \frac{\pi_1^{\mathrm{T}}\mathbf{T}_{11}\mathbf{e}}{\pi_1^{\mathrm{T}}\mathbf{e}}\right)^{-1}$ and we get

$$\mathbf{W}_2 = \mathbf{r}\alpha\mathbf{e}^{\mathrm{T}}, \quad \mathbf{W}_{a2} = \mathbf{r}\alpha_a\mathbf{e}^{\mathrm{T}}.$$

Now we verify that $\alpha = \alpha_a$. Firstly we evaluate α_a

$$\alpha_a = \mathbf{v}_1^{\mathrm{T}}\mathbf{e}\left(1 - \frac{\pi_1^{\mathrm{T}}\mathbf{T}_{11}\mathbf{e}}{\pi_1^{\mathrm{T}}\mathbf{e}}\right)^{-1} = \mathbf{v}_1^{\mathrm{T}}\mathbf{e}\left(\frac{\pi_1^{\mathrm{T}}\mathbf{e} - \pi_1^{\mathrm{T}}\mathbf{T}_{11}\mathbf{e}}{\pi_1^{\mathrm{T}}\mathbf{e}}\right)^{-1} = \frac{\mathbf{v}_1^{\mathrm{T}}\mathbf{e}\;\pi_1^{\mathrm{T}}\mathbf{e}}{\pi_1^{\mathrm{T}}\,(\mathbf{I} - \mathbf{T}_{11})\,\mathbf{e}}.$$

From Lemma 2 we get

$$\pi_2^{\mathrm{T}}\mathbf{T}_{21} = \pi_1^{\mathrm{T}}\,(\mathbf{I} - \mathbf{T}_{11}) \tag{30}$$

and then

$$\alpha_a = \frac{\mathbf{v}_1^{\mathrm{T}}\mathbf{e}\;\pi_1^{\mathrm{T}}\mathbf{e}}{\pi_2^{\mathrm{T}}\mathbf{T}_{21}\mathbf{e}} = \frac{\mathbf{v}_1^{\mathrm{T}}\mathbf{e}\;\pi_1^{\mathrm{T}}\mathbf{e}}{\pi_2^{\mathrm{T}}\mathbf{u}_2\mathbf{v}_1^{\mathrm{T}}\mathbf{e}} = \frac{\pi_1^{\mathrm{T}}\mathbf{e}}{\pi_2^{\mathrm{T}}\mathbf{u}_2}.$$

Now we evaluate α

$$\alpha \;=\; \mathbf{v}_1^{\mathrm{T}}\left(\mathbf{I}-\mathbf{T}_{11}\right)^{-1}\mathbf{e}\,.$$

From the relation (30) we have

$$\begin{aligned}\pi_2^{\mathrm{T}}\mathbf{u}_2\mathbf{v}_1^{\mathrm{T}} &= \pi_1^{\mathrm{T}}\left(\mathbf{I}-\mathbf{T}_{11}\right)\\ \mathbf{v}_1^{\mathrm{T}} &= \frac{1}{\pi_2^{\mathrm{T}}\mathbf{u}_2}\pi_1^{\mathrm{T}}\left(\mathbf{I}-\mathbf{T}_{11}\right)\end{aligned}$$

and substitute it to the formula for α. We get

$$\alpha=\frac{1}{\pi_2^{\mathrm{T}}\mathbf{u}_2}\pi_1^{\mathrm{T}}\left(\mathbf{I}-\mathbf{T}_{11}\right)\left(\mathbf{I}-\mathbf{T}_{11}\right)^{-1}\mathbf{e}=\frac{\pi_1^{\mathrm{T}}\mathbf{e}}{\pi_2^{\mathrm{T}}\mathbf{u}_2}=\alpha_a.$$

Thus $\mathbf{M}_{22}=\mathbf{M}_{a22}$.

Remark 1 *If* $\mathbf{T}_{21}=\mathbf{u}_2\mathbf{v}_1^{\mathrm{T}}$, $\mathbf{T}_{12}=\mathbf{u}_1\mathbf{v}_2^{\mathrm{T}}$, *where* $\mathbf{u}_1=(0,\ldots,0,1)^{\mathrm{T}}$, $\mathbf{u}_2=(1,0,\ldots,0)^{\mathrm{T}}$, $\mathbf{v}_1=(0,\ldots,0,t_{n_1+1,n_1})^{\mathrm{T}}$, $\mathbf{v}_2=(t_{n_1,n_1+1},0,\ldots,0)^{\mathrm{T}}$ *then the BM property transformes into the cut point situation.*

The mean first passage matrix to the aggregated matrix to SPV π is

$$\mathbf{M}_{[\mathbf{A}(\pi)]^{\mathrm{T}}}=\begin{pmatrix}8.8873 & 3.5500 & 5.0555 & 18.0777 & 17.0777 & 16.0777\\ 28.000 & 3.1550 & 1.5055 & 14.5277 & 13.5277 & 12.5277\\ 30.000 & 2.0000 & 2.8044 & 13.0222 & 12.0222 & 11.0222\\ 31.000 & 3.0000 & 1.0000 & 14.0222 & 13.0222 & 12.0222\\ 32.000 & 4.0000 & 2.0000 & 1.0000 & 14.0222 & 13.0222\\ 33.000 & 5.0000 & 3.0000 & 2.0000 & 1.0000 & 14.0222\end{pmatrix} \tag{31}$$

3.4 Aggregations for Column Computation

For arbitrary $k-$th column: we solve the equation

$$\left[\mathbf{I}-\left(\mathbf{T}-\mathbf{T}_{\mathrm{k}}\right)\right]\mathbf{M}_{\mathrm{k}}=\mathbf{e}, \tag{32}$$

where $\mathbf{T}_{\mathrm{k}}$, $\mathbf{M}_{\mathrm{k}}$ are the $k-$th columns of the corresponding matrices.

We apply the iteration process:

$$\begin{aligned}\mathbf{M}_k^{(0)} &= \mathbf{e},\\ \mathbf{M}_k^{(i+1)} &= \left(\mathbf{T}-\mathbf{T}_{*k}\right)\mathbf{M}_k^{(i)}+\mathbf{e}.\end{aligned} \tag{33}$$

Since $\rho(\mathbf{T}-\mathbf{T}_{*k})<1$, then the iteration process converges, but the convergence is slow. There is possible to apply the following aggregation process to speed up the computation.

Algorithm: $\widetilde{\mathbf{C}}=[\mathbf{R}\left(\mathbf{I}-\left(\mathbf{T}-\mathbf{T}_{*k}\right)\right)^{\mathrm{T}}\mathbf{S}(\pi)]^{\mathrm{T}}$=$[\mathbf{R}\,\mathbf{A}_i^{\mathrm{T}}\,\mathbf{S}(\pi)]^{\mathrm{T}}$ (see (23), (24))

For $i=1,\ldots,$ consider the residuals: $\mathbf{r}_i=\mathbf{e}-[(\mathbf{I}-(\mathbf{T}-\mathbf{T}_{*k}))]\,\mathbf{M}_k^{(i)}$,

$\widetilde{\mathbf{C}}\,\mathbf{o}_i=[\mathbf{S}(\pi)]^{\mathrm{T}}\,\mathbf{r_i}$

$\mathbf{M}_k^{(i+\frac{1}{2})}=\mathbf{M}_k^{(i)}+\mathbf{R}^{\mathrm{T}}\,\mathbf{o_i}$

$\widetilde{\mathbf{M}}^{(0)}=\mathbf{M}^{(i+\frac{1}{2})}$

$\widetilde{\mathbf{M}}^{(j+1)}=\left(\mathbf{T}-\mathbf{T}_{*k}\right)\widetilde{\mathbf{M}}_k^{(j)}+\mathbf{e}$, $j=1,\ldots,s$

$\mathbf{M}^{(i+1)}=\widetilde{\mathbf{M}}^{(s)}$.

The convergence of the algorithm is assured. We can state

Theorem 4 *There exists such s, that the algorithm converges.*

The proof is clear from the text.

4 Cost Analysis

In this paper an overview of methods for computing the mean first passage times matrix is given. The method introduced by C. D. Meyer (see [9]) requires approximately $4/3\,n^3$ operations and the inversion of a large full matrix. Another access (shown in [3]) needs approximately $7/6\,n^3$ operations and requires two matrix inversions of matrices of size n/2. The access presented by P. Mayer requires not more than n^3 operations and requires the inversion of a matrix (but in case when the transition matrix is a sparse then the inverted matrix is a sparse as well and the number of operations needed can be further reduced).

A common drawback of all approaches is the necessity to compute entire matrix what is not effective in particular when we are interested in a few of its elements and all the elements of the matrix are useless. The processes presented in this paper eliminate this deficiency.

5 Conclusion

The possibilities of computation of a part of MFPTM using aggregation procedures are introduced in this paper. This approach eliminates the need of inversion in particular of the block $\mathbf{I} - \mathbf{T}_{11}$ which significantly reduces severity of computational processes. The stationary probability vector (SPV) is necessary to know but this vector can be effectively obtained using the IAD algorithm presented here.

Reference

[1] KEMENY J. G., SNELL J. L.: *Finite Markov Chains*, Van Nostrand, New York, 1960.

[2] KIRKLAND S. J., NEUMANN M.: Cutpoint decoupling and first passage times for randomwalks ongraphs, *SIAM J. Matrix Anal. Appl.*, 20(4): 860–870.

[3] KIRKLAND S. J., NEUMANN M., Xu J.: A divide and conquer approach to computing the mean first passage matrix for Markov chains via Perron complement reductions, *Numer. Linear Algebra Appl.*, 8: 287–295, 2001.

[4] KLAPKA Š., MAYER P.: Some aspects of modeling railway safety, In *Proceedings of SANM'1999*, pp. 135–140, Plzeň, 1999, University of West Bohemia.

[5] KLAPKA Š., MAYER P.: Využití matematického modelování při koncepčním řešení předmětných úkolů, *Technical Report 1*, AŽD Praha s. r. o., 2000, Internal report, in Czech.

[6] KLAPKA Š., MAYER P.: Aggregation/disaggregation method for safety models, *Applications of Mathematics*, 47(2), pp. 127–137, 2002.

[7] MAREK I., MAYER P.: Iterative aggregation/disaggregation methods for computing some characteristics of Markov chains, *Large Scale Scientific Computing, Third International Conference, LSSC 2001,*, pp. 68–82, Sozopol, Bulgaria, 2001.

[8] MAYER P.: Numerical methods for Markov chain models, *Engineering Mechanics*, Vol 13, No. 3, pp. 163–176, 2006.

[9] MEYER C. D. J.: The role of the group generalized inverse in the theory of finite Markov chains, *SIAM Review*, 17:443–464, 1975.

[10] MEYER C. D. J.: Uncoupling the Perron eigenvector problem, *Linear Algebra and Its Applications 1989*, 114/115:69–94.

[11] STEWART W. J.: *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, 1994.

Optimization of Linear Differential Systems with Delay by Lyapunov's Direct Method

H. Demchenko, J. Diblík, D. Khusainov

Brno University of Technology, Brno University of Technology, Taras Shevchenko National University of Kyiv
xdemch02@stud.feec.vutbr.cz, diblik@feec.vutbr.cz, khusainov@unicyb.kiev.ua

Abstract: *Direct Lyapunov's method is applied to solve optimization problems for linear differential equations with delay. Optimal control functions minimizing quality criteria are found.*

Keywords: Lyapunov function, differential system with delay, optimal control function, quality criterion.

Introduction

As it is well-known, there are two approaches to solving optimization problems in dynamic systems. The first approach was proposed by L.S. Pontryagin. His method is based on finding a fixed control (a program control) for which the solution of the system described by differential equations reaches a predetermined previous value and minimizes the integral quality criterion. The second approach consists in finding a control function in the form of a feedback such that the trivial solution is asymptotically stable and simultaneously minimize the integral quality criterion. Being based on the second Lyapunov method, this is a kind of a dynamic programming method (a combination of methods in calculus of variations and Lyapunov functions method). Its founder is N.N. Krasovskii. We refer, e.g., to [1, 3, 4, 5]. In this paper, the latter method is applied to linear differential systems with delay.

1 Problem considered

Consider a control system described by a system of differential equations with delay

$$\frac{dx(t)}{dt} = f\left(x(t), x(t-\tau), u(t,x)\right), \tag{1}$$

where $f\colon R^n \times R^n \times R^m \to R^n$, $f = (f_1, \ldots, f_n)$, $f\left(0,0,0\right) = 0$, $(x, v, u) \in \mathfrak{D}$,

$$\mathfrak{D} := \{(x, v, u) \in R^n \times R^n \times R^m, t \geq t_0\},$$

$t_0 \in R$, $\tau > 0$, $n \geq 1$, $m \geq 1$ are natural numbers, and the undisturbed system

$$\frac{dx(t)}{dt} = f\left(x(t), x(t-\tau), 0\right). \tag{2}$$

We need to find a control function in the form of a feedback, i.e., $u = u(t, x)$, such that a solution $x(t)$, $t \geq t_0$ of system (1) corresponding to this control and to the initial function is asymptotically

stable and the integral (often called quality criterion)

$$I = \int_{t_0}^{\infty} \omega\left(x(t), x(t-\tau), u(t,x)\right) dt \tag{3}$$

attains a minimum value. The function $\omega(x, y, u)$ defined on $\mathfrak{D}$ is assumed to be positive definite.

Let $C_\tau^n = C([-\tau, 0), \mathbb{R}^n)$ be the space of the continuous mappings from the interval $[-\tau, 0)$ into $\mathbb{R}^n$. If A is any set in $\mathbb{R}^n$, we will set $C_\tau^n(A) = C([-\tau, 0), A)$.

Let $C_\tau^n(D)$ be the space of the continuous mappings from the interval $[-\tau, 0)$ into the set $D = \{\xi \in \mathbb{R}^n \colon \|\xi\| < M\}$, M is a positive constant (or $M = \infty$).

Let $x \colon [t_0 - \tau, \infty) \to \mathbb{R}^n$ be a continuous vector-function, $t_0 \in \mathbb{R}$, and let $\tau > 0$ be a number. For a given $t \in [t_0, \infty)$, we define a norm

$$\|x(t)\|_\tau = \max_{\theta \in [-\tau, 0]} (\|x(t+\theta)\|)$$

where

$$\|x(s)\| = \max_{i=1,\ldots,n} \{|x_i(s)|\},\ s \in [t_0 - \tau, \infty).$$

The following three definitions and Theorem 1 are taken from [2]

Definition 1 *Let a functional $V \colon (\alpha, \infty) \times C_\tau^n(D) \to \mathbb{R}$ be given. It is called positive-definite if there exists a continuous nondecreasing function $w \colon [0, M) \to \mathbb{R}$, $w(0) = 0$, $w(s) > 0$ if $s \in [0, M)$ such that*

$$V(t, \psi) \geq w(\|\psi(0)\|)$$

on $(\alpha, \infty) \times C_\tau^n(D)$.

Definition 2 *Let a functional $V \colon (\alpha, \infty) \times C_\tau^n(D) \to \mathbb{R}$ be given. V is said to have an infinitesimal upper bound if there exists a continuous nondecreasing function $W \colon [0, M) \to \mathbb{R}$, $W(0) = 0$, $W(s) > 0$ if $s \in [0, M)$ such that*

$$V(t, \psi) \leq W(\|\psi\|_\tau)$$

on $(\alpha, \infty) \times C_\tau^n(D)$.

Definition 3 *A positive-definite functional $V \colon (\alpha, \infty) \times C_\tau^n(D) \to \mathbb{R}$ having an infinitesimal upper bound is called a Lyapunov-Krasovskii functional.*

Theorem 1 *If there exists a Lyapunov-Krasovskii functional*

$$V \colon (\alpha, \infty) \times C_\tau^n(D) \to \mathbb{R}$$

and if $V(t, x_t)$ defines a nonincreasing function of t on $[t_0, \beta)$ whenever

$$x = x(\cdot, t_0, \varphi),\ t \in [t_0 - \tau, \beta)$$

is the noncontinuable solution of (2) through some $(t_0, \varphi) \in (\alpha, \infty) \times C_\tau^n(D)$, then the trivial solution of (2) is asymptotically stable.

Define an auxiliary function

$$B\left(V,t,x(t),x_t,u\right) := \frac{dV(t,x_t)}{dt} + \omega(x(t),x(t-\tau),u) \tag{4}$$

where V is a Lyapunov-Krasovskii functional and $dV(t,x_t)/dt$ denotes the derivative of V with respect to t along trajectories of system (1). The following theorem was motivated by a similar theorem for non delayed systems [4].

Theorem 2 *Assume that, for the system of differential equations (1), there exists a Lyapunov-Krasovskii functional $V_0(t,x_t)$ having an infinitesimal upper bound and a function $u_0(t,x)$ such that*
1. The function $\omega(x(t),x(t-\tau),u_0(t,x))$ is positive-definite for every $t \geq t_0$, $\|x\| < M$, where M is a positive constant.
2. $B(V_0,t,x(t),x_t,u_0(t,x)) \equiv 0$.
3. $B(V_0,t,x(t),x_t,u(t,x)) \geq 0$ for any $u(t,x) \not\equiv u_0(t,x)$.

Then, the function $u_0(t,x)$ is a solution of the optimal stabilization problem and

$$\int_{t_0}^{\infty} \omega(x(t),x(t-\tau),u_0(t,x))dt$$
$$= \min_u \left[\int_{t_0}^{\infty} \omega(x(t),x(t-\tau),u(t,x))dt\right] = V_0(t_0,x_{t_0}). \tag{5}$$

Proof. The functional $V_0(t,x_t)$ satisfies all conditions of Theorem 1. For its derivative along trajectories of the system (1), we have

$$\frac{dV_0}{dt} = -\omega(x(t),x(t-\tau),u_0(t,x)), \tag{6}$$

which means that it is a negative-definite function. That is why, for $u = u_0(t,x)$, the undisturbed motion $x(t) \equiv 0$ is asymptotically stable and $\lim_{t\to\infty} x(t) = 0$ for all initial conditions $x(t_0)$ from the region $\|x(t_0)\|_\tau \leq \eta$, where η can be found from the equation

$$\sup[V_0(t,x_t)|_{\|x\|_\tau \leq \eta}] < \inf[V_0(t,x_t)|_{\|x\|_\tau = h}],$$

and $h < M$.

Now it is sufficient to show that (5) is true. The motion $x_0(t)$ satisfies condition $\|x_0(t)\|_\tau \leq h < M$. Thus, for all $t \geq t_0$, the equation (6) holds. Moreover, from the property of asymptotic stability, we have

$$\lim_{t\to\infty} V_0(t,x_{0t}) = 0. \tag{7}$$

Integrating equation (6) along the motion $x_0(t)$ over (t_0,∞), using (7), we obtain

$$V_0(t_0,x_{t_0}) = \int_{t_0}^{\infty} \omega(x_0(t),x_0(t-\tau),u_0(t,x))dt. \tag{8}$$

On the other hand, let $u = u_*(t,x)$ be an arbitrary function that is also a solution of the optimal stabilization problem for the motion $x(t) \equiv 0$ and for initial conditions $\|x(t_0)\|_\tau \leq \eta$. Assume that, for $t \geq t_0 - \tau$, $x_*(t)$ lies inside the region $\|x(t)\|_\tau \leq h$. Then, by assumption 3, we get

$$\frac{dV_0}{dt} \geq -\omega(x_*(t),x_*(t-\tau),u_*(t,x)). \tag{9}$$

Integrating this inequality over (t_0, ∞) and using the property

$$\lim_{t\to\infty} V_0(t, x_{*t}) = 0 \tag{10}$$

we obtain

$$V_0(t_0, x_{t_0}) \le \int_{t_0}^{\infty} \omega(x_*(t), x_*(t-\tau), u_*(t, x))dt. \tag{11}$$

A similar inequality can be obtained if the motion $x_*(t)$ goes out of the region $\|x(t)\|_\tau \le h$ on an interval. In this case, we have the following situation. Let $t_1 > t_0$ be the moment of time, when the motion $x_*(t)$ goes back into the region and stays in it for all $t \ge t_1$. Then, from that moment on, equation (9) will hold for $x_*(t)$. Integrating this inequality over (t_1, ∞) and using the property (10) again, we obtain

$$V_0(t_1, x_{*t_1}) \le \int_{t_1}^{\infty} \omega(x_*(t), x_*(t-\tau), u_*(t, x))dt. \tag{12}$$

Since $x(t_0)$ satisfies $\|x(t_0)\|_\tau \le \eta$, where η is sufficiently small, we have

$$V_0(t_0, x_{t_0}) < V_0(t_1, x_*(t_1)), \tag{13}$$

and, due to assumption 1, we get

$$\int_{t_1}^{\infty} \omega(x_*(t), x_*(t-\tau), u_*(t, x))dt < \int_{t_0}^{\infty} \omega(x_*(t), x_*(t-\tau), u_*(t, x))dt. \tag{14}$$

From (12)–(14), we derive (11), and from (8), (11) we get (5). □

2 Linear equations

Consider linear scalar equations with constant coefficients and a single delay

$$\frac{dx(t)}{dt} = ax(t) + bx(t-\tau) + cu(x(t)), \tag{15}$$

where a, b, c are real constants, $\tau > 0$ is a delay and $u(x(t))$ is a control function.

Together with equation (15), we will consider a quality criterion (3) with $t_0 = 0$ and

$$\omega(x(t), x(t-\tau), u) = \alpha x^2(t) + 2\beta x(t)x(t-\tau) + \gamma x^2(t-\tau) + \delta u^2(x(t)),$$

i.e., (3) being a quadratic criterion

$$I = \int_0^{\infty} \left[\alpha x^2(t) + 2\beta x(t)x(t-\tau) + \gamma x^2(t-\tau) + \delta u^2(x(t))\right] dt, \tag{16}$$

with $\alpha > 0$, $\alpha\gamma - \beta^2 > 0$, $\delta > 0$.

Theorem 3 *If*

$$\beta b < 0 \tag{17}$$

and

$$b(\alpha + \gamma) = 2a\beta, \tag{18}$$

then the optimal stabilization control function

$$u_0(x(t)) = 2\frac{\beta c}{b\delta}x(t) \tag{19}$$

exists.

Proof. We utilize Theorem 2. Define a Lyapunov-Krasovskii functional

$$V(t,x_t)=hx^2(t)+\int_{t-\tau}^{t}dx^2(s)ds,\ h>0,\ d>0. \tag{20}$$

Then, in accordance with condition 2 of Theorem 2, we analyze the expression B given by (4), i.e.,

$$B\left(V,t,x(t),x_t,u\right)=2hx(t)[ax(t)+bx(t-\tau)+cu(x(t))]+$$

$$+d[x^2(t)-x^2(t-\tau)]+\alpha x^2(t)+2\beta x(t)x(t-\tau)+\gamma x^2(t-\tau)+\delta u^2(x(t))=0.$$

Simplifying the last expression, we get

$$B\left(V,t,x(t),x_t,u\right)=(2ha+d+\alpha)x^2(t)+(\gamma-d)x^2(t-\tau)+$$

$$+(2hb+2\beta)x(t)x(t-\tau)+2hcx(t)u(x(t))+\delta u^2(x(t))=0.$$

This equation will be satisfied if

$$2ha+d=-\alpha, \tag{21}$$

$$d=\gamma, \tag{22}$$

$$hb=-\beta, \tag{23}$$

$$2hcx(t)u(x(t))=-\delta u^2(x(t)). \tag{24}$$

Condition (23) is valid because of (17). Substituting (22) and (23) into (21), we get condition (18).

From (24), we obtain an optimal control in the form

$$u_0(x(t))=-2\frac{hc}{\delta}x(t)=2\frac{\beta c}{b\delta}x(t).$$

□

Example 1 *Consider equation (15) with* $a=-2$, $b=-1$, $c=1$, *i.e.,*

$$\dot{x}(t)=-2x(t)-x(t-\tau)+u(t)$$

with a quadratic quality criterion (16) with $\alpha=2>0$, $\beta=1$, $\gamma=2$, $\delta=1>0$, *(here* $\alpha\gamma-\beta^2=3>0$*), i.e.,*

$$I=\int_0^{\infty}(2x^2(t)+2x(t)x(t-\tau)+2x^2(t-\tau)+u^2(t))dt.$$

Since $\beta b=-1<0$ *and* $b(\alpha+\gamma)=2a\beta=-4$, *all assumptions of Theorem 3 are true. By formula (19), the optimal stabilization control function*

$$u_0(x(t))=2\frac{\beta c}{b\delta}x(t)=-2x(t)$$

exists.

3 Linear systems

Consider linear systems with constant coefficients with one constant delay

$$\frac{dx(t)}{dt} = A_0 x(t) + A_1 x(t-\tau) + bu(x(t)), \tag{25}$$

where A_0, A_1 are $n \times n$ constant matrices, $b \in R^n$, $u(x(t)) \in R$, and a quality criterion (3) with $t_0 = 0$ and

$$\omega(x(t), x(t-\tau), u(t,x)) := x^T(t) C_{11} x(t) + x^T(t) C_{12} x(t-\tau) +$$

$$+ x^T(t-\tau) C_{21} x(t) + x^T(t-\tau) C_{22} x(t-\tau) + du^2(x(t)),$$

where $C_{11}, C_{12}, C_{21}, C_{22}$ are $n \times n$ positive-definite matrices, C_{11} and C_{22} are symmetric, $C_{21} = C_{12}^T$ and $d > 0$, i.e., (3) is a quadratic criterion

$$I = \int_0^\infty \left[x^T(t) C_{11} x(t) + x^T(t) C_{12} x(t-\tau) + \right.$$

$$\left. + x^T(t-\tau) C_{21} x(t) + x^T(t-\tau) C_{22} x(t-\tau) + du^2(x(t)) \right] dt. \tag{26}$$

Theorem 4 *Assume that there exists a positive definite symmetric matrix H satisfying Lyapunov matrix equation*

$$A_0^T H + H A_0 = -C_{11} - C_{22}. \tag{27}$$

If, moreover,

$$A_1^T H = -C_{21}, \tag{28}$$

the optimal stabilization control function

$$u_0(x(t)) = -\frac{2}{d} b^T H x(t) \tag{29}$$

exists.

Proof. We utilize Theorem 2. Define a Lyapunov-Krasovskii functional

$$V(t, x_t) = x^T(t) H x(t) + \int_{t-\tau}^{t} x^T(s) G x(s) ds,$$

where H, G are $n \times n$ positive-definite matrices. Then, in accordance with condition 2 of Theorem 2, we analyze the expression B given by (4), i.e.,

$$\begin{aligned} B\left(V, t, x(t), x_t, u\right) = & \left[A_0 x(t) + A_1 x(t-\tau) + bu(x(t))\right]^T H x(t) \\ & + x^T(t) H \left[A_0 x(t) + A_1 x(t-\tau) + bu(x(t))\right] + \\ & + x^T(t) G x(t) - x^T(t-\tau) G x(t-\tau) \\ & + x^T(t) C_{11} x(t) + x^T(t) C_{12} x(t-\tau) + x^T(t-\tau) C_{21} x(t) \\ & + x^T(t-\tau) C_{22} x(t-\tau) + du^2(x(t)) \equiv 0. \end{aligned}$$

Simplifying the last expression, we get

$$B\left(V,t,x(t),x_t,u\right) = x^T(t)[A_0^T H + HA_0 + G + C_{11}]x(t) + \\ + x^T(t-\tau)[A_1^T H + C_{21}]x(t) + x^T(t)[HA_1 + C_{12}]x(t-\tau) + \\ + x^T(t-\tau)[C_{22} - G]x(t-\tau) + \\ + b^T u(x(t))Hx(t) + x^T(t)Hbu(x(t)) + du^2(x(t)) \equiv 0.$$

This will hold if

$$A_0^T H + HA_0 = -C_{11} - G, \tag{30}$$

$$A_1^T H = -C_{21}, \tag{31}$$

$$HA_1 = -C_{12}, \tag{32}$$

$$G = C_{22}, \tag{33}$$

$$u(x(t))[b^T Hx(t) + x^T(t)Hb] = -du^2(x(t)). \tag{34}$$

Equation (30) is valid due to (33) and (27). Equations (31) and (32) hold due to (28).

From (34), we obtain an optimal control in the form

$$u_0(x(t)) = -\frac{2}{d}b^T Hx(t).$$

□

Example 2 *Consider system (25) with* $n = 2$ *and*

$$A_0 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \; A_1 = \begin{pmatrix} -1/2 & \varepsilon \\ 0 & -1/2 \end{pmatrix}, \; b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

where ε *is an arbitrary constant, i.e.,*

$$\dot{x_1}(t) = -x_1(t) - \frac{1}{2}x_1(t-\tau) + \varepsilon x_2(t-\tau) + b_1 u(t),$$
$$\dot{x_2}(t) = -x_2(t) - \frac{1}{2}x_2(t-\tau) + b_2 u(t)$$

with a quadratic quality criterion (26) with

$$C_{11} = \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix}, \; C_{12} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}, \; C_{21} = \begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix}, \; C_{22} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

i.e.,

$$I = \int_0^\infty \Bigg[(x_1(t), x_2(t)) \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + (x_1(t), x_2(t)) \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} x_1(t-\tau) \\ x_2(t-\tau) \end{pmatrix} \\ + (x_1(t-\tau), x_2(t-\tau)) \begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \\ + (x_1(t-\tau), x_2(t-\tau)) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1(t-\tau) \\ x_2(t-\tau) \end{pmatrix} + du^2(t) \Bigg] dt.$$

We assume that δ is a constant such that

$$|\delta| < 1 \tag{35}$$

We show that all assumptions of Theorem 4 are true. From equation (27), we get

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} + \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = - \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Simplifying, we obtain

$$H = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} = \begin{pmatrix} 1 & \delta/2 \\ \delta/2 & 1 \end{pmatrix}.$$

Matrix H is a positive definite symmetric matrix if $|\delta| < 2$. This inequality is valid due to (35). The matrix C_{11} is positive definite, too. Consider equation (28). In our case, we get

$$\begin{pmatrix} -1/2 & 0 \\ \varepsilon & -1/2 \end{pmatrix} \begin{pmatrix} 1 & \delta/2 \\ \delta/2 & 1 \end{pmatrix} = - \begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix}$$

and

$$C_{21} = \begin{pmatrix} 1/2 & \delta/4 \\ \delta/4 - \varepsilon & 1/2 - \varepsilon\delta/2 \end{pmatrix}.$$

Since $C_{12} = C_{21}^T$, we get

$$C_{12} = \begin{pmatrix} 1/2 & \delta/4 - \varepsilon \\ \delta/4 & 1/2 - \varepsilon\delta/2 \end{pmatrix}.$$

Obviously, C_{12}, C_{21} are positive definite matrices for an arbitrary ε. The matrix C_{22} is positive definite as well. So, all assumptions are fulfilled and Theorem 4 is applicable. By formula (29), the optimal stabilization control function

$$u_0(x(t)) = -\frac{2}{d} b^T H x(t) = -\frac{2}{d} (b_1, b_2) \begin{pmatrix} 1 & \delta/2 \\ \delta/2 & 1 \end{pmatrix} x(t)$$

exists.

4 Conclusion

The paper applies a method developed by N.N. Krasovskii to solving optimal stabilization problems for differential equations and systems with delay. This method makes it possible to find a control function in the form of a feedback such that the zero solution of a given equation or system is asymptotically stable and, simultaneously, an integral quality criterion attains a minimum value.

Acknowledgements

The first two authors were supported by the Grant FEKT-S-14-2200 of Faculty of Electrical Engineering and Communication, Brno University of Technology. The third author has been supported by the Development Programme of Ministry of Education, Youth and Sports of the Czech Republic No. 3.5 - Support of international mobility of academic staff of Brno University of Technology.

Reference

[1] ALEKSEEV, V.M., TIKHOMIROV, V.M., FOMIN, S.V.: *Optimal Control.* Moskow, Nauka, 1979. 432 pp.

[2] DRIVER, R.D.: *Ordinary and Delay Differential Equations.* Springer-Verlag New York Inc., 1977.

[3] GANTMAKHER, F.R.: *Theory of Matrices.* Moskow, Nauka, 1966. 576 pp.

[4] MALKIN, I.G.: *Theory of Stability of Motion.* Moskow, Nauka, 1966. 530 pp.

[5] PONTRYAGIN, L.S., BOLTYANSKII, V.G., GAMKRILEDZE, R.V., MISCHENKO, E.F.: *Mathematical Theory of Optimal Processes.* Moskow, Nauka, 1983. 392 pp.

Stability and controllability of treatments models and security

Irada Dzhalladova

Department of Computer Mathematics and Information Security,
V. Getman Kyev National Economic University,
Kyev, Ukraine,
idzhalladova@gmail.com

Abstract: *We study controllability and stability properties of dynamical systems when actuator or sensor signals are under attack. We formulate a detailed adversary model that considers different levels of privilege for the attacker such as read and write access to information flows. We then study the impact of these attacks and propose reactive countermeasures based on game theory.*

The primary line of defense for any system is its proactive security mechanisms. Therefore, in practice we must use the threat model to identify the most valuable targets for an adversary and invest in protecting them.

We consider open-loop solutions. To find the necessary conditions for optimality of the situation we need to use Pontryagins minimum principle.

Keywords: stability,optimality, security of dynamical systems, actuator or sensor signals, attack, information war.

1 Introduction

We study controllability and stability properties of dynamical systems when actuator or sensor signals are under attack. We formulate a detailed adversary model that considers different levels of privilege for the attacker such as read and write accesss to information flows. We then study the impact of these attacks and propose reactive countermeasures based on game theory.

The security of cyber-physical control systems has received significant attention in the last couple of years [1,2]. The primary line of defense for any system are its proactive security mechanisms. Therefore, in practice we must use the threat model to identify the most valuable targest for an adversary and invest in protecting them.

If an attack is detected, the defender can respond with different actions. Some of the possible responses include reconfiguration of the system, attack isolation, of even a system shutdown (for safety reasons). In this work we are interested in defenses that respond to attacks by changes in their control actions; thus creating a game-theory problem where the actions on the players are the control signals each of them has assess to. In particular, we assume that if the system is not under attack, the system will operate with a vanilla control signal $u(t)$; however, when the system detects an attack, in chages to a reactive control signal $u_s(t)$ to maintain the system under the best possible conditions. This creates a differential game between the defender and the attacker. We use a recept model for data integrity attacks in demand-response programs for the smart grid [3]. The model considers actuator attacks as an aggregate effect for multi-agent systems that all receive the same input control signal.

2 Statement of problem

May be the most general framework in control system is the theory of Linear Time Invariant state space system [4]. In this setting we consider a system of linear differential equations

$$\dot{x(t)} = Ax(t) + Bu(t) \tag{1}$$

$$y(t) = Cx(t) + Du(t) \tag{2}$$

where $x(t) \in R^n$ is a vector of phisical quantities representing the state of the system at time t, $u(t) \in R^p$ is the control input at time t, $y(t) \in R^n$ is a vector of sensor measuarements at time t, and A,B, C, D - matrics representing the dynamics of the system.

2.1 Control and security properties

Similar to security properties such as confidentiality, integrity, and availiability, there are several control properties that asystem designer or plant operators would like to maintain,even under attack. In the theory of linear state space systems, two dual properties are controllability and observability.

Controllability is an important property of a control system, and the controllability property plays a crucial role in many control problems, such as stabilization of unstable systems by feedback, or optimal control. Controllability means that the state of the system can be driven to any arbitrary place by using the manipulated variables.

2.2 Attacking Controlability

We define an attack model for control systems containig three part: goal of the attacker,offline information,and online information. While in general setting an attacker can have many different objectives, in this paper we focus on attackers that try to manipulate the controllability or stability of the system. Using the attacker model , we turn to problem of how controllability and stability can be attacked. This anaysis can be used for risk assessment by identifying the resiliency of the system to attacks or to identify the actuators and sensors that are most valuable to the system.

Let us consider one of the interesting and general case system (1) (Attacking Controlability with $u(t)$): the linear control system

$$\frac{dX(t)}{dt} = A(t, \xi(t))X(t) + B(t, \xi(t))U(t) \tag{3}$$

on the probability basis $\left(\Omega, \mathfrak{T}, \mathbf{P}, \mathbf{F} \equiv \{\mathbf{F_t} : t \geq 0\}\right)$ and together with (3) we consider the initial conditions

$$X(0) = \varphi(\omega), \qquad \varphi : \Omega \to \mathbb{R}^n.$$

The coefficients of the system are semi-Markov coefficients defined by the transition intensities $q_{\alpha k}(t), \alpha, k = 1, 2, \cdots, n$, 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 $q_{\alpha k}(t), \alpha, k = 1, 2, \cdots, n$, satisfy the conditions [8]:

$$q_{\alpha k}(t) \geq 0, \qquad \int_0^\infty q_k(t)\, dt = 1, \qquad q_k(t) \equiv \sum_{\alpha=1}^{\infty} q_{\alpha k}(t).$$

If $\psi_k(t)$ denotes the probability of the event that no jump takes place during the interval (t_j, t_{j+1}), provided that the process jumps to the state θ_k at time t_j, then

$$\psi_k(t) = \int_t^\infty q_k(\tau)d\tau, \qquad k = 1, 2, \cdots, n. \tag{4}$$

In our considerations, it will be convenient to denote the block-diagonal matrix,

$$\Psi(t) = \operatorname{diag}\big(\psi_1(t), \psi_2(t), \cdots, \psi_n(t)\big). \tag{5}$$

Definition 1. *Let the matrices $Q(t, \xi(t))$, $L(t, \xi(t))$ with semi-Markov elements be symmetric and positive definite. The cost functional*

$$J = \int_0^\infty \Big\langle X^*(t)Q(t, \xi(t))X(t) + U^*(t)L(t, \xi(t))U(t)\Big\rangle\, dt, \tag{6}$$

defined on the space $C^1 \times U$, where $\langle\cdot\rangle$ denotes mathematical expectation, is called ***the quality criterion****.*

Definition 2. *Let $S(t, \xi(t))$ be a matrix with semi-Markov elements. The control vector*

$$U(t) = S(t, \xi(t))X(t) \tag{7}$$

which minimizes the quality criterion $J(X, U)$ with respect to the system (3) is called ***the optimal control****.*

If we denote

$$G(t, \xi(t)) \equiv A(t, \xi(t)) + B(t, \xi(t))\, S(t, \xi(t)),$$
$$H(t, \zeta(t)) \equiv Q(t, \zeta(t)) + S^*(t, \zeta(t))\, L(t, \zeta(t))\, S(t, \zeta(t)),$$

then the system (3) can be rewritten to the form

$$\frac{dX(t)}{dt} = G(t, \xi(t))X(t) \tag{8}$$

and the functional (6) to the form

$$J = \int_0^\infty \Big\langle X^*(t)H(t, \xi(t))X(t)\Big\rangle\, dt. \tag{9}$$

We suppose also, that together with every jump of random process $\xi(t)$ in time t_j, the solutions of the system (8) submit to the random transformation

$$X(t_j + 0) = C_{sk}X(t_j - 0), \qquad s, k = 1, 2, \cdots n,$$

if the conditions $\xi(t_j + 0) = \theta_s$, $\xi(t_j - 0) = \theta_k$ hold.

Definition 3. *Let $a_k(t)$, $k = 1, \cdots, n$, $t \geq 0$ be a selection of n different positive functions. If $\xi(t_j + 0) = \theta_s$, $\xi(t_j - 0) = \theta_k$, $s, k = 1, \cdots n$, 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.

2.3 Stability properties

Another impornant property of control system is a stability. Several different stability definitions are useful. Here, we recall the mean stability and the mean square stability definitions, the L_2 stability given in [5], and the classical definition of asymptotic stability.

Definition 4. *The trivial solution of system (3) is said to be mean square stable on the interval* $[0,\infty)$ *if for each* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that any solution* $X(t)$ *corresponding to the initial data* $X(0)$ *exists for all* $t \geq 0$ *and the mathematical expectation*

$$E^{(1)}\{\|X(t)\|^2\} < \varepsilon \quad \textit{whenever} \quad t \geq 0 \quad \textit{and} \quad \|X(0)\| < \delta.$$

The mean stability of the zero solution of system (3) is defined in much the same way with only $\|X(t)\|^2$ being replaced by $\|X(t)\|$.

Several different stability definitions are useful. Here, we recall the mean stability and the mean square stability definitions, the L_2 stability given in [5,7], and the classical definition of asymptotic stability.

Definition 5. *The trivial solution of system (3) is said to be mean square stable on the interval* $[0,\infty)$ *if for each* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that any solution* $X(t)$ *corresponding to the initial data* $X(0)$ *exists for all* $t \geq 0$ *and the mathematical expectation*

$$E^{(1)}\{\|X(t)\|^2\} < \varepsilon \quad \textit{whenever} \quad t \geq 0 \quad \textit{and} \quad \|X(0)\| < \delta.$$

The mean stability of the zero solution of system (3) is defined in much the same way with only $\|X(t)\|^2$ being replaced by $\|X(t)\|$.

Definition 6. *The trivial solution of the differential systems (3) is said to be* L_2 *stable if the integral*

$$\int_0^\infty E^{(1)}\{\|X(t)\|^2\}dt \tag{10}$$

converges.

3 Main results

The optimal control $U(t)$ for the system (3) 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 (3) have special properties or intensities $q_{sk}(t)$ satisfy some relations or some other special conditions are satisfied [10,11,12]. Some of these cases will be formulated as corollaries.

Theorem 1. *Let the control system* (3) *with piecewise constant coefficients have the form*

$$\frac{dX(t)}{dt} = A(\xi(t))X(t) + B(\xi(t))U(t). \tag{11}$$

Then the quadratics functional

$$V = \int_0^\infty \left\langle X^*(t)Q(\xi(t))X(t) + U^*(t)L(\xi(t))U(t) \right\rangle dt \tag{12}$$

determines the optimal control in the form

$$U(t) = S(t, \xi(t))X(t),$$

where

$$S(t, \xi(t)) = S_k(t - t_j)$$

and the matrices $S_k(t)$ satisfy the equations

$$S_k(t) = -L^{-1}B_k^* R_k(t), \quad k = 1, 2, \cdots, n \tag{13}$$

if $t_j \le t < t_{j+1}$, $\xi(t) = \theta_k$.
The matrices $R_k(t),\ k = 1, 2, \cdots, n$ are the solutions of the systems of the Riccati type equations:

$$\begin{aligned}\frac{dR_k(t)}{dt} = &-Q_k - A_k^* R_k(t) - R_k(t)A_k \\ &+ R_k(t)B_k L_k^{-1} B_k^* R_k(t) - \frac{\Psi_k'(t)}{\Psi_k(t)} R_k(t) \\ &- \sum_{s=1}^{n} \frac{q_{sk}(t)}{\Psi_k(t)} C_{sk}^* R_s(0) C_{sk}, \quad k = 1, \cdots, n.\end{aligned} \tag{14}$$

Remark 1. *In the corollary we mention piecewise constant coefficients of the control system (11). The coefficients of the functional (12) will be piecewise as well, but the optimal control is unstationary.*

Corollary 1. *Let us assume that*

$$\frac{\Psi_k'(t)}{\Psi_k(t)} = \text{const}, \quad \frac{q_{sk}(t)}{\Psi_k(t)} = \text{const}, \quad k, s = 1, 2, \cdots, n. \tag{15}$$

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 $R_k(t),\ k = 1, 2, \cdots, n$ in (13) are constant, which implies the form of the system (14) is changed to the form

$$\begin{aligned}&Q_k + A_k^* R_k + R_k A_k - R_k B_k L_k^{-1} B_k^* R_k + \frac{\Psi_k'(t)}{\Psi_k(t)} R_k(t) \\ &+ \sum_{s=1}^{n} \frac{q_{sk}(t)}{\Psi_k(t)} C_{sk}^* R_k C_{sk} = 0, \quad k = 1, \cdots, n.\end{aligned} \tag{16}$$

The system (16) has constant solutions $R_k,\ k = 1, 2, \cdots, n$, if conditions (15) hold. Moreover, if the random process $\xi(t)$ is a Markov process then the conditions (15) have the form

$$\frac{\Psi_k'(t)}{\Psi_k(t)} = a_{kk} = \text{const}, \quad \frac{q_{sk}(t)}{\Psi_k(t)} = a_{sk} = \text{const}, \quad k, s = 1, 2, \cdots, n,\ k \ne s,$$

and the system (16) transforms to the form

$$Q_k + A_k^* R_k + R_k A_k - R_k B_k L_k^{-1} B_k^* R_k + \sum_{s=1}^{n} a_{sk} C_{sk}^* R_s C_{sk} = 0,\ k = 1, \cdots, n$$

for which the optimal control is

$$U(t) = S(\xi(t))X(t),\ \ S(\theta_k) \equiv S_k,\ \ S_k = -L_k^{-1} B_k^* R_k,\ \ k = 1, 2, \cdots, n.$$

Corollary 2. *Let the state θ_s of the semi-Markov process $\xi(t)$ is not be longer than $T_s > 0$. Then the system (12) has the form*

$$\begin{aligned} \nu_k(x) &\equiv x^* C_k x \\ &= \int_0^{T_s} \Bigg(X_k^*(t) \Big(\Psi_k(t) Q_k(t) + \sum_{s=1}^{n} q_{sk}(t) C_{sk}^* C_s C_{sk} \Big) X_k(t) \\ &\qquad + U_k^*(t) \Psi_k(t) L_k(t) U_k(t) \Bigg) dt, \qquad k = 1, 2, \cdots, n. \end{aligned} \tag{17}$$

Because

$$K_s(T_s) = \Psi_s(t) R_s(t),\ \ s = 1, 2, \cdots, n,$$

then

$$K_s(T_s) = 0,\ \ s = 1, 2, \cdots, n. \tag{18}$$

In this case, the search for the matrix $K_s(t),\ s = 1, 2, \cdots, n$ in concrete tasks is reduced to integration of the matrix system of differential equations (6) on the interval $[0, T_s]$ with initial conditions (18). In view of $\Psi_s(T_s) = 0,\ s = 1, 2, \cdots, n$, we can expect, that every equation (11) has a singular point $t = T_s$. If $\Psi_s(t)$ has simple zero at the point $t = T_s$, then the system (12) meets the necessary condition

$$\Psi_s(T_s) R_s(T_s) + \sum_{k=1}^{n} q_{sk}(T_s) C_{ks}^* R_s(0) C_{ks} = 0,\ \ s = 1, \cdots, n$$

for boundary of matrix $R_s(t)$ in the singular points.

4 Model problem

Let the semi-Markov process $\xi(t)$ take two states $\theta_1,\ \theta_2$ and let it be identical with the Markov process described by the system of differential equations

$$\begin{aligned} \frac{dp_1(t)}{dt} &= -\lambda p_1(t) + \lambda p_2(t), \\ \frac{dp_2(t)}{dt} &= \ \ \lambda p_1(t) - \lambda p_2(t). \end{aligned}$$

We will consider the L_2-stability of the solutions of the differential equation

$$\frac{dx(t)}{dt} = a\big(\xi(t)\big)x(t), \quad a(\theta_k) \equiv a_k \tag{19}$$

constructing a system of the type (14) related to the equation (19). The system is

$$c_1 = 1 + \int_0^\infty e^{2a_2 t}\lambda e^{-\lambda t} c_2\, dt, \qquad c_2 = 1 + \int_0^\infty e^{2a_1 t}\lambda e^{-\lambda t} c_1\, dt$$

and its solution is

$$c_1 = \frac{(\lambda - a_1)(\lambda - 2a_2)}{2a_1 a_2 - \lambda(a_1 + a_2)}, \qquad c_2 = \frac{(\lambda - a_2)(\lambda - 2a_1)}{2a_1 a_2 - \lambda(a_1 + a_2)}.$$

The trivial solution of the equation (19) is L_2-stable, if $c_1 > 0$ and $c_2 > 0$. Let the intensities of semi- Markov process $\xi(t)$ satisfy the conditions

$$q_{11}(t) \approx 0, \;\; q_{22}(t) \approx 0, \;\; q_{21}(t) - \lambda e^{-\lambda t} \approx 0, \;\; q_{12}(t) - \lambda e^{-\lambda t} \approx 0$$

Then, using the Theorem 1, the conditions

$$1 - c_1 \int_0^\infty q_{11}(t) e^{2a_1 t} dt - c_2 \int_0^\infty \left(q_{21}(t) - \lambda e^{-\lambda t}\right) e^{2a_2 t} dt > 0,$$
$$1 - c_1 \int_0^\infty \left(q_{12}(t) - \lambda e^{-\lambda t}\right) e^{2a_1 t} dt - c_2 \int_0^\infty q_{22}(t) e^{2a_2 t} dt > 0.$$

are sufficient conditions for the L_2-stability of solutions of the equation (19).

Threat modeling is a procedure for optimizing network security by identifying objectives and vulnerabilities, and then defining countermeasures to prevent, or mitigate the effects of, threats to the system. In this context, a threat is a potential or actual adverse event that may be malicious (such as a denial-of-service attack) or incidental, and that can compromise the assets of an enterprise.

Security threat modeling, or threat modeling, is a process of assessing and documenting a system's security risks. Security threat modeling enables our to understand a system's threat profile by examining it through the eyes of our potential foes. With techniques such as entry point identification, privilege boundaries and threat trees, you can identify strategies to mitigate potential threats to your system. Our security threat modeling efforts also enable your team to justify security features within a system, or security practices for using the system, to protect our corporate assets.

There are some aspects to security threat modeling(with example in economics situation):

1. **Identify threats**. *For example,* our system's ordering module interacts with the payment processing module. Anybody can place an order, but only manager-level employees can credit a customer's account when he or she returns a product. At the boundary between the two modules, someone could use functionality within the order module to obtain an illicit credit.

2. **Understand the threat(s).** To understand the potential threats at an entry point, you must identify any security-critical activities that occur and imagine what an adversary might do to attack or misuse our system.

On questions such as "How could the adversary use an asset to modify control of the system, retrieve restricted information, manipulate information within the system, cause the system to fail or be unusable, or gain additional rights.

In this way, we can determine the chances of the adversary accessing the asset without being audited, skipping any access control checks, or appearing to be another user. To understand

the threat posed by the interface between the order and payment processing modules, we would identify and then work through potential security scenarios.
For example, an adversary who makes a purchase using a stolen credit card and then tries to get either a cash refund or a refund to another card when he returns the purchase.

3. **Categorize the threats.** To categorize security threats, consider the STRIDE (Spoofing, Tampering, Repudiation, Information disclosure, Denial of Service, and Elevation of privilege) approach. Classifying a threat is the first step toward effective mitigation.
For example, if we know that there is a risk that someone could order products from our company but then repudiate receiving the shipment, we should ensure that you accurately identify the purchaser and then log all critical events during the delivery process.

4. **Identify mitigation strategies.** To determine how to mitigate a threat, we can create a diagram called a threat tree. At the root of the tree is the threat itself, and its children (or leaves) are the conditions that must be true for the adversary to realize that threat. Conditions may in turn have subconditions.
For example, under the condition that an adversary makes an illicit payment. The fact that the person uses a stolen credit card or a stolen debit/check card is a subcondition. For each of the leaf conditions, we must identify potential mitigation strategies; in this case, to verify the credit card using the some verification package and the debit card with the issuing financial institution itself. Every path through the threat tree that does not end in a mitigation strategy is a system vulnerability.

5. **Test**. Our threat model becomes a plan for penetration testing. Penetration testing investigates threats by directly attacking a system, in an informed or uninformed manner.
Informed penetration tests are effectively white-box tests that reflect knowledge of the system's internal desig, whereas uninformed tests are black box in nature.

References

[1] AMIN, S., SCHWARTZ, G.A., Shankar Sastry.: Security of interdependent and identical networked control systems. Automatica (2012) .

[2] PASQUALETTI, F.: Secure control systems: A control-theoretic approach to cyber-physical security. Ph.D. dissertation, University of California (2012) .

[3] ASTROM K. J., Introduction to Stochastic Control Theory, *New York, Academic Press*, (1970)

[4] CARLOS B.,A.Cardenes, Nicanor Quijano. Controllabillity of Dynamics Systems:treat models and reactive Security. 4th International Conference, GameSec 2013 Fort Worth, TX, USA, November 2013 Proceedings. pp. 45-65.

[5] DZHALLADOVA, I, RUZICHKOVA, M: *The Optimization of Solutions of the Dynamic Systems with Random Structure*, Abstract and Applied Analysis, Hindawi Publishing Corporation, Vol. 2011, pp. 1-18, 2011. ISSN: 1085-3375.

[6] VALEEV, K G, DZHALLADOVA, I: *Optimization of Nonlinear Systems of Stochastic Difference Equations*, Ukrainian Mathematical Journal, Vol. 54, No. 1, pp. 1-16, 2002. ISSN: 1573-9376.

[7] DIBLIK, J, DZHALLADOVA, I, RUZICHKOVA, M: *The Stability of Nonlinear Differential Systems with Random Parameters*, Abstract and Applied Analysis, Hindawi Publishing Corporation, Vol. 2012, pp. 1-11 , 2012. ISSN: 1085-3375

[8] VALEEV, K G, DZHALLADOVA, I: *Optimization of a System of Linear Differential Equations with Random Coefficients*, Ukrainian Mathematical Journal, Vol. 51, No. 2, pp. 622-629, 1999. ISSN: 1573-9376.

[9] BAŠTINEC, J.; DZHALLADOVA, I.:*Sufficient conditions for stability of solutions of systems of nonlinear differential equations with right-hand side depending on Markov's process.* In 7. konference o matematice a fyzice na vysokých školách technických s mezinárodní účastí. 2011. pp. 23-29. ISBN 978-80-7231-815-5.

[10] DZHALLADOVA, I. Stability of the ecosystems models of global processes. MITAV, Brno, 2015. pp. 3-11.

[11] DIBLÍK, J., KHUSAINOV, D.Y., BAŠTINEC, J., RYVOLOVÁ, A.: *Exponential stability and estimation of solutions of linear differential systems with constant delay of neutral type.* In 6. konference o matematice a fyzice na vysokých školách technických s mezinárodní účastí. Brno, UNOB Brno. 2009. pp. 139-146. ISBN 978-80-7231-667-0.

[12] DZHALLADOVA, I.; BAŠTINEC, J.; DIBLÍK, J.; KHUSAINOV, D.: *Estimates of exponential stability for solutions of stochastic control systems with delay.* Abstract and Applied Analysis. 2011. 2011(1). pp. 1-14. ISSN 1085-3375. (IF=1,318).

INTERACTIVE SCHOOL EXPERIMENTS IN THE PSE GRAPHICAL ENVIRONMENT

Fabo Peter, Pavlíková Soňa

Research Centre, University of Žilina
Univerzitná 8215/1, 010 26 Žilina, Slovakia
fabo.peter@gmail.com

Faculty of Chemical and Food Technology, Slovak University of Technology in Bratislava
Radlinského 9, 812 37 Bratislava, Slovakia
sona.pavlikova@stuba.sk

Abstract: *The paper presents usage of PSE graphical environment based on the Python programming language for creating simple educational experiments. PSE environment allows solving problems using graphical components - a high level blocks with defined properties as well as a creation of user-defined components. With the help of simple and readily available hardware components such as the Arduino, it is possible creating demonstration experiments. In the second part of the paper we will show some possibilities for the demonstration of solutions of linear and nonlinear differential equations with examples of classic bifurcation diagrams. By way of simple examples, are shown basic characteristics of Z-transform and its use in the implementation of digital filters.*

Keywords: programming, education, python, simulation, numpy, scipy, matplotlib, difference equations

INTRODUCTION

PSE (Python Simulator Editor) is an open-source block-oriented simulation environment developed in Python, primarily focused on the creation of general simulation models. The environment uses the extensive infrastructure of Python [1], PyQt application framework [2], libraries for scientific computing NumPy and SciPy [3] and visualization library Matplotlib [4]. In the development of PSE environment was placed a major emphasis on its openness, the user can modify the environment, expand and add new components, and thus it differs from other commercial and open alternatives.

The environment allows you creating simulation models in the form of diagrams by the transfer of mathematical relations into visual form through components that contain individual algorithms to transform information and oriented connections between individual components in Fig. 1. Connections in each simulation step move data among components and the information can be scalar or vector.

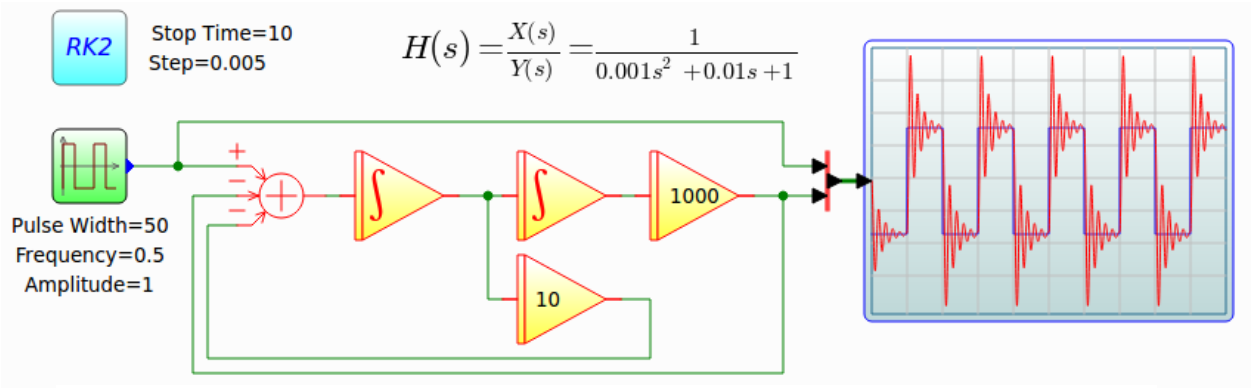


Fig. 1. A typical diagram of the PSE environment. The implementation of the transfer function and the time domain response to an input signal, simulation algorithm is Runge-Kutta method of the second order.

The components are arranged in libraries, classed according to their functionality or meaning. Typical library components include:

- Sources - sources of information - generators, data from a file, retrieve information from connected devices and the Internet
- Sinks – information consumers- write to a file, console output, graphs, send information into the Internet
- Control – components for communication control
- Linear – components for linear transformation of information
- Nonlinear – components for non-linear transformation of information
- Signal – components for editing connections, aggregation of scalar connections to vector and vector connections to scalar
- Discrete – discrete and logic components
- Interactive – components for interactive control of the diagram during a simulation

Components of all the groups in the diagram can be combined freely. Individual library components for better orientation in the diagram graphically distinct. The simulation of larger diagrams is possible by creating separate diagrams – blocks, and use them in the simulation as separate components. The blocks can be used as separate components as they are expanded in the diagram as a macro with a separate namespace, Fig. 2., Fig. 3.

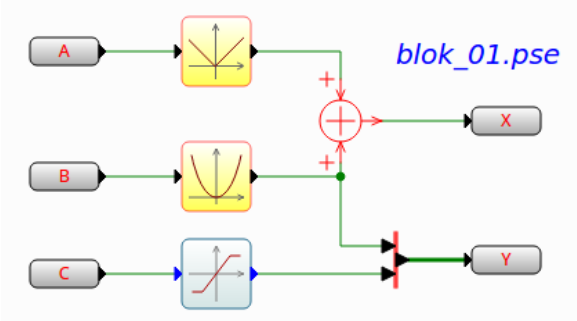


Fig. 2. The diagram used to form the block. It does not contain the simulation control component

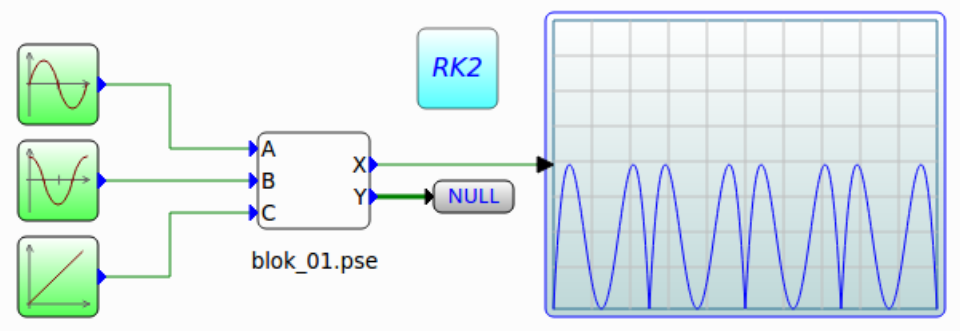


Fig. 3. Use of the block from Fig. 2. in a simulation. The block is identified by the name of the diagram from which was created and it is possible to use it repeatedly

1. ELEMENTARY INTERACTIVE DEMONSTRATION IN THE PSE ENVIRONMENT

There are used only the properties of the PSE environment, user interaction with a simulated process if necessary, is conducted through any of the standard user interface components (Button, Slider, Dial, etc.) inserted into the diagram. The purpose of the simulation is the demonstration of the topic with the possibility of changing the parameters of the simulated process, Fig. 4., Fig. 5..

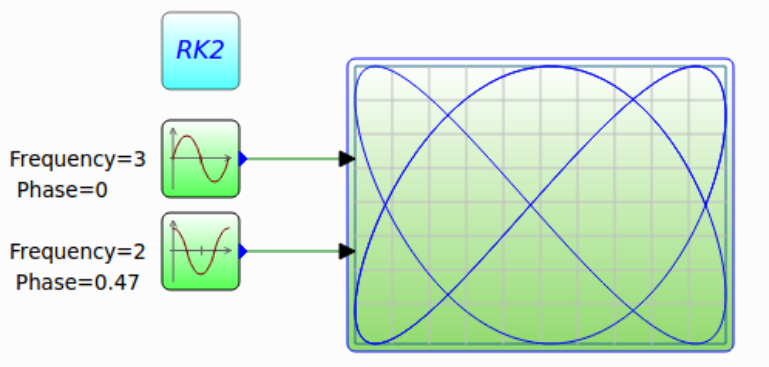


Fig. 4. Elemental demonstration of perpendicular oscillations composition. Simulation results can be varied by changing the operating parameters of the signal components.

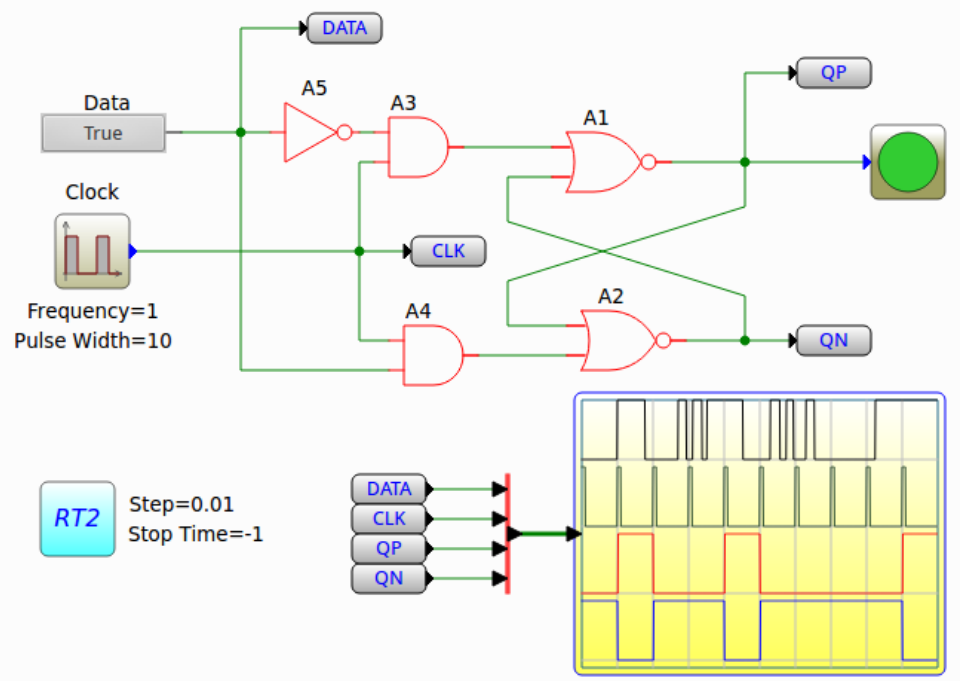


Fig. 5. Simulation of a logic circuit, gated RS toggle circuit in real time (StopTime = -1). Data on the input can be changed using the True/False button, status signals are displayed in the real time on a graph and the status of data output is shown by a 'LED' indicator.

2. SIMPLE EXPERIMENTS WITH TECHBOARD INPUT MODULE

For teaching the fundamentals of programming using the Scratch [5] application, was developed TechBoard input module. The module is backwards compatible with the original PicoBoard [6] module, but with wider options of peripherals connectivity and robust mechanical construction. The module communicates via USB and communication in Python is possible using the standard library Serial.

The standard module PicoBoard contains light sensor, sound sensor, button, slide potentiometer input and four (AD) inputs for measuring voltage resistive divider. TechBoard as a modified version, allows you connecting external devices - joystick, buttons, light barriers, resistance temperature sensor, etc. Inputs A and C are numerically linearized, so that when connected potentiometer, the output value is linearly proportional to the resistance. The values of the inputs B and D are proportional to the voltage on resistive divider and the top resistor of the divider (2k) is a part of the module. In the PSE environment is the module represented by the block, which generates a vector of the sensor and input values. Application example of module usage is shown in Fig. 6.

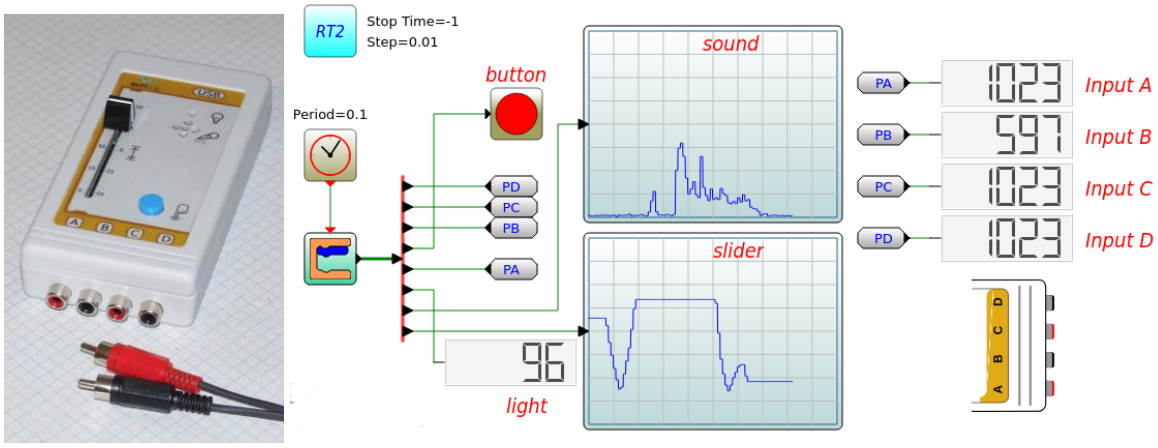


Fig. 6. TechBoard module and its usage in the PSE environment

The module allows interactive control of simulation, like the components of the PSE environment. More interesting in educational practice is the use the module for simple physical experiments in elementary and secondary schools, especially in the field of time data collection from sensors – i.e. demonstration of exponential course of the cooling liquid, light conditions during the day and the passing clouds. Period of data collection module can be controlled by a timer, the minimum period is 0.1 sec.

3. ARDUINO AS AN PSE INTERFACE TO THE REAL WORLD

Arduino [7] is a popular and affordable platform for teaching the microcontrollers technology, control and robotics. The Arduino programming environment is available with a library for connecting typical peripherals - motors, servos, interface I2C, SPI, and many others. From the pedagogical point of view Arduino suffers from (and also all similar platform based on a microcontroller) a problem of tracking and debugging the program. Control algorithm is stored in the memory of the microcontroller and without special aids it is possible to learn about the state of a program only through a change in the state of selected module terminals or over UART to the console.

Using a simple programs downloaded to the Arduino allows its usage as a relatively quick input-output device connected via USB, while at the maximum communication speed is real time of information exchange between the PSE and the Arduino platform less than 2 msec, which is sufficient for regular school experiments. Since the Arduino platform is rather flexible, its inputs and outputs can be optimized for the given experiment. From the pedagogical point of view is important, however, that the actual control of the experiment runs under the PSE environment in real time. It is therefore possible to interactively optimize and present the state variables of a controlled process. A simple use example is the optical target tracking - infrared diodes using two reception diodes, rotated by servos, Fig. 7., Fig. 8., and video [8].

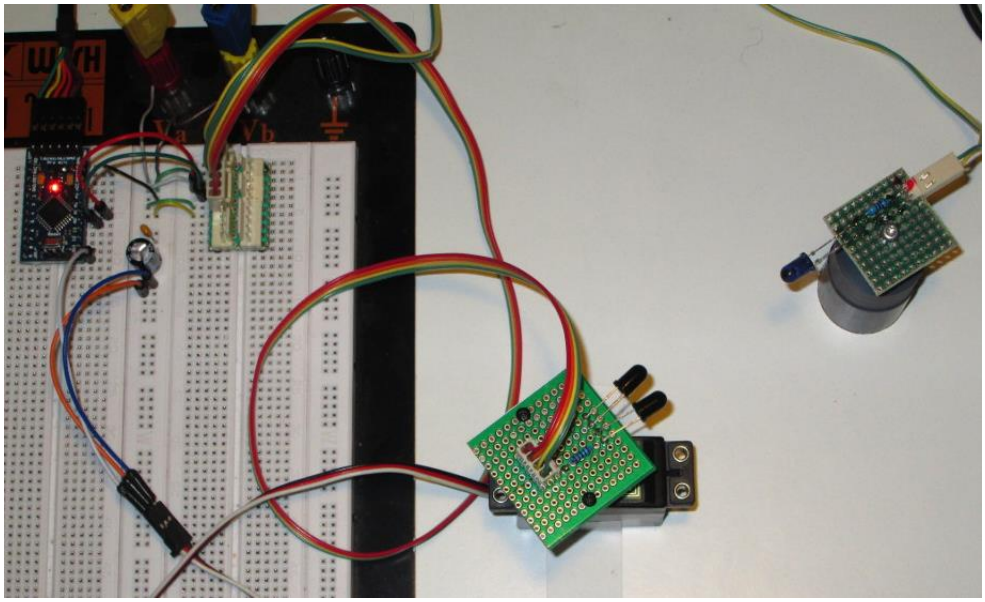


Fig. 7. The experiment set with objective of target tracking (optical tracking system), servo is controlled by Arduino PWM output, two receiving photodiodes are connected to the inputs of A/D converters of Arduino A0 and A1.

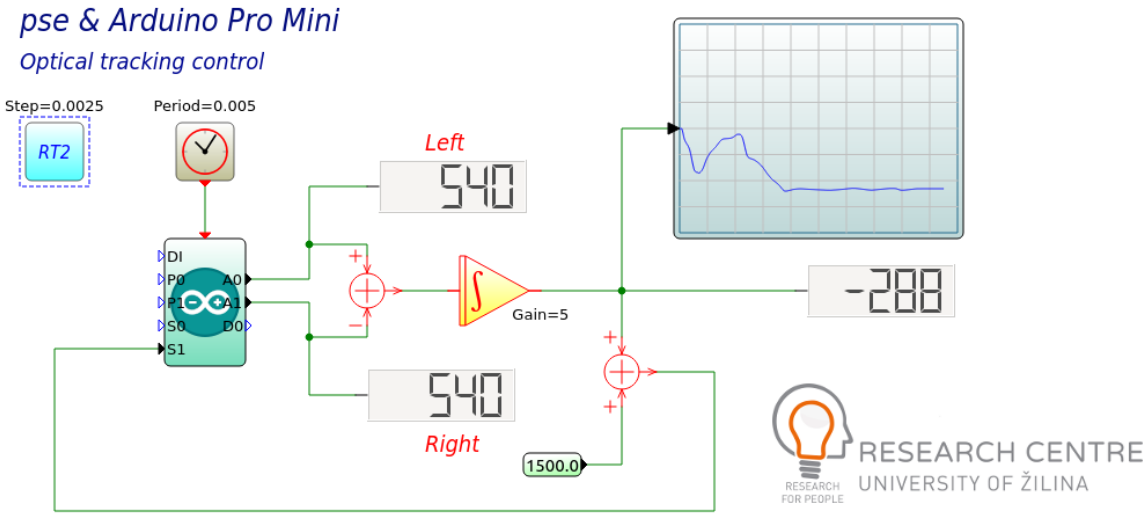


Fig. 8. Control of the experiment from Fig. 7. Communication with Arduino is represented by a component that mediates communication and a conversion of input-output values.

4. THE USE OF THE PSE BLOCK GRAPHICS USER INTERFACE FOR A DEMONSTRATION OF SOLVING DIFFERENCE EQUATIONS

We will present features of open source PSE graphical environment, based on the Python programming language and its possibilities for the demonstration of solutions of differential linear and nonlinear equations with examples of bifurcation diagrams. The essential characteristics of Z-transform and its use in the implementation of digital filters will be shown through simple examples.

Differential count has an important application in Mathematics itself (in numerical mathematics, especially in the numerical solution of differential equations, probability theory, number theory), but also in such applications as in civil and electrical engineering.

For *n = 0, 1, ... , n* and function *x(n)* is defined difference operator Δ as follows:

$$\Delta x(n) = x(n+1) - x(n)$$

Higher ordinary differences for natural number m are defined as follows:

$$\Delta^{n+1} x(n) = \Delta\left[\Delta^n x(n)\right]$$

Difference equation of one independent variable $n \in N$ and one unknown function $u(n)$ is a functional equation that has the form:

$$f\left(n, u(n), u(n+1), ..., u(n+k)\right) = 0$$

Let us show some simple examples of difference equations and systems of difference equations.

Examples:

1. Logistic map

$$x(n+1) = rx(n)\left(1 - x(n)\right)$$

r is a given constant, x_0 is the initial value and each sequence is determined by the given equation.

2. Digital filter

Digital filter with finite impulse response (FIR) can be described by difference equation of the form

$$y_n = \sum_{k=0}^{n-1} h_k x_{n-k}$$

The digital filter with infinite impulse response (IIR) is characterized by recursive difference equation in the form

$$\sum_{m=0}^{M-1} a_m y_{n-m} = \sum_{k=0}^{n-1} b_k x_{n-k}$$

3. Predator-prey model is described by a system of difference equations

$$x(n+1) - x(n) = -ax(n) + bx(n)y(n), a, b > 0$$
$$y(n+1) - y(n) = py(n) - qx(n)y(n), p, q > 0$$

and set values *x(0), y(0).*

4. Model rivarly

$$x(n+1) - x(n) = ax(n) - bx(n)y(n), a, b > 0$$
$$y(n+1) - y(n) = py(n) - qx(n)y(n), \ \text{p}, q > 0$$

Difference equations can be solved in addition to analytical methods using Z - transform. In the following examples are presented solutions of difference equations using

simulation models. Examples are useful in education especially for students of study fields in informatics, automation, robotics, electrical and so on.

PSE simulation environment was created as an open source software, developed exclusively in Python [1] using its extensive libraries. The emphasis in the development of PSE was first put on its openness, possibilities of expansion and modifications, which is different from the commercial option.

PSE was primarily developed for the creation of simulation models, its usage is also as an interactive tool for demonstration of selected topics in the teaching process at secondary schools and universities.

5. DIFFERENCE EQUATIONS SIMULATIONS

A typical difference equation consists of a combination of input and output values mutually displaced in the multiples of discrete time intervals. The basic component of the simulation of difference equations is therefore a component time unit shift of the input value in the standard Z-transform marked as z^{-1}.

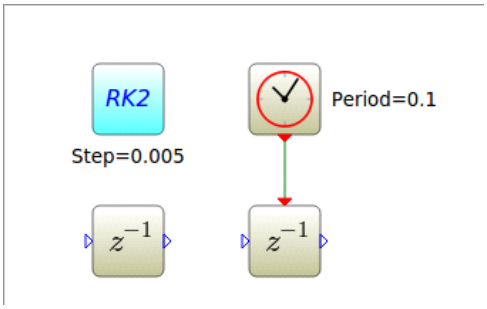


Fig. 9. Component unit delay. The delay time is defined by Solvera iteration step and the delay time is defined by an external timing.

Examples

Fig. 10. shows a block structure with a simple difference equation representing the band digital filter of the second-order. The use of the block to filter a periodic signal is shown in Fig. 11. The time delay is defined by the simulation parameter.

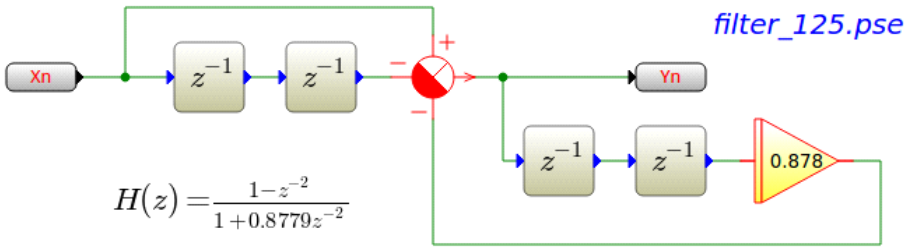


Fig. 10. IIR block structure with difference equation.

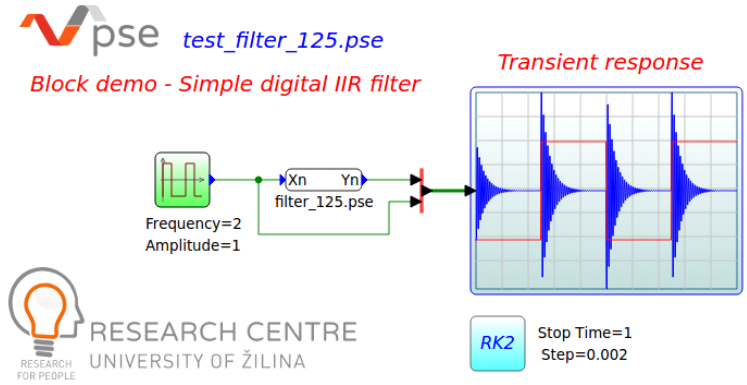


Fig. 11. Usage ob block from Fig. 10.

Block structure of a more complicated difference equation is in Fig. 12.

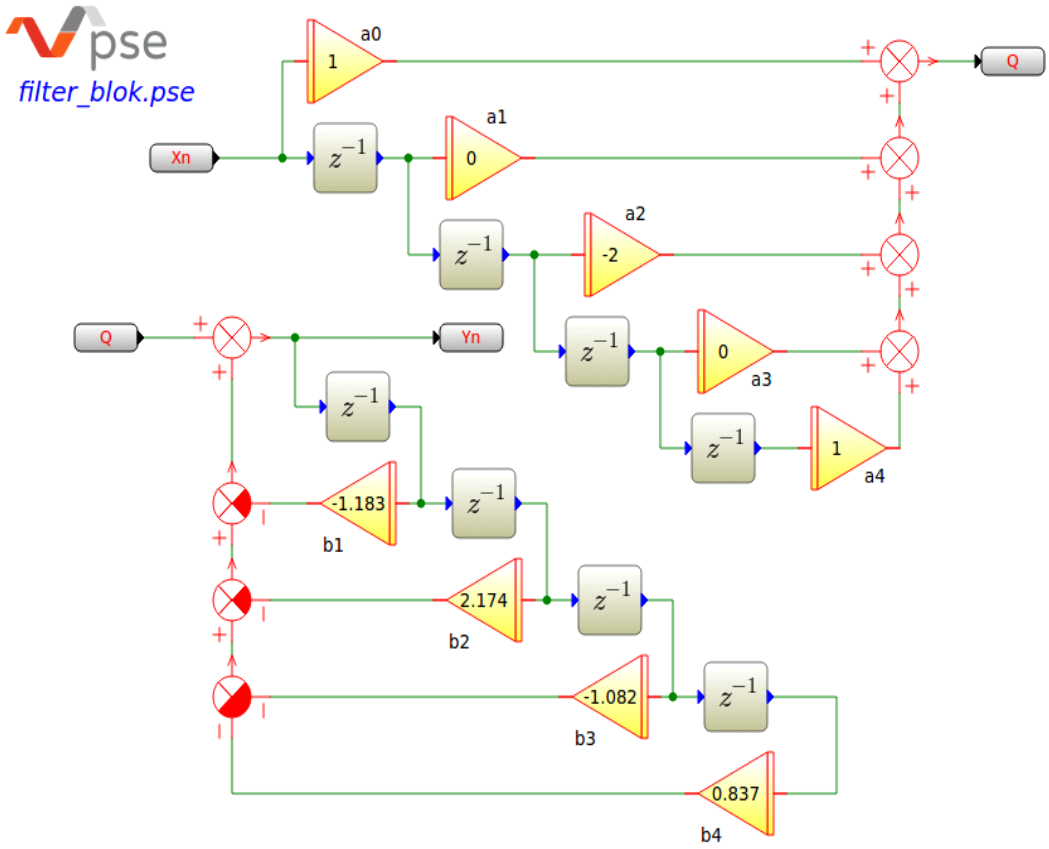


Fig. 12. The structure of differential equation of IIR digital filter divided into an input and output section.

Difference equations describing the evolution of the population in the time, referred to as logistical view, can produce non-stationary solutions and chaos. Non stationary parametric solution simulation of difference equations is shown in Fig. 13.

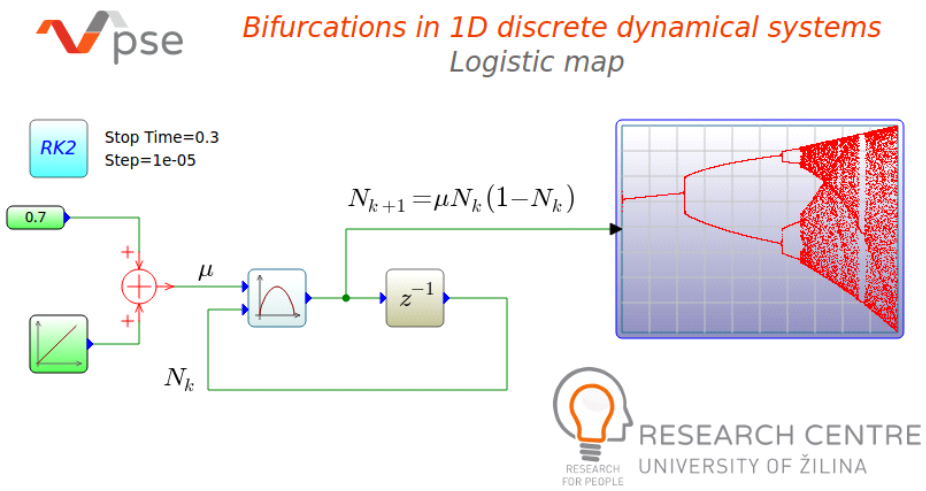


Fig. 13. Bifurcations in one-dimensional discrete dynamical system.

CONCLUSION

The limited scope of this paper does not describe all the possibilities of the PSE platform in the pedagogical process in details. In short, we can mention the creation of interactive textbooks in the environment Ipython Notebook [9], management of laboratory equipment via TCP/IP or specialized buses and acquisition respectively processing data from the experiments. Attractive is also a communication of separate PSE environments in the Internet environment via UDP packets and the ability to create interactive experiments distributed to students, such as observation of the weather in a wider geographical area, sharing of joint experiments and the like. At the stage of experimental verification are specialized input-output modules for applications in robotics, mechatronics and physics experiments.

Platform PSE is freely available at [10], documentation, examples, and tutorials are available at [11].

REFERENCES

[1] https://www.python.org/

[2] http://www.riverbankcomputing.co.uk/software/pyqt/intro

[3] Jones E, Oliphant E, Peterson P, et al. SciPy: *Open Source Scientific Tools for Python*, 2001-, http://www.scipy.org/ [Online; accessed 2015-04-29].

[4] Hunter, J. D., Matplotlib: *A 2D graphics environment*, Computing In Science & Engineering, vol. 9., number 3, pages 90-95 (2007).

[5] https://scratch.mit.edu/

[6] https://www.sparkfun.com/products/11888

[7] http://www.arduino.cc/

[8] https://youtu.be/_8YgjFWc0S0

[9] http://ipython.org/notebook.html

[10] https://158.193.150.40:2443/public/projects

[11] https://pse.fri.uniza.sk/

[12] REKTORYS, K*: Přehled užité matematiky*, SNTL - Nakladatelství technické literatury, Praha 1981

Acknowledgement
pse Framework was developed primarily for the creation of simulation models for predictive simulation parameters of road infrastructure within the project at Research Centre of the University of Žilina.

The research is supported by the European Regional Development Fund and the Slovak state budget by the projects "Research Centre of the University of Žilina" - ITMS 26220220183.

RAISING ATTRACTIVENESS IN TEACHING TECHNICAL SUBJECTS BY USING SOFTWARE MEANS

Erika Fechová
Faculty of Manufacturing Technologies with a seat in Prešov,
Department of Mathematics, Informatics and Cybernetics, Technical University of Košice
Bayerova 1, 080 01 Prešov, Slovakia,
erika.fechova@tuke.sk

Abstract: *Introduction of modern information and communication means is significant means of modernizing and raising attractiveness in educational process. The main goal of these modernizing tendencies is continual increase of educational level corresponding to the current level of knowledge in individual subjects with a respect to the age of a student and type of a school, which one studies at. Information and communication means also became the part of university education, where their utilization mainly serves to development of inter-subject relations. The paper deals with the possibility of using application software such as Matlab, MS Excel and Mathematica in educational process of natural science and technical subjects. Suitability of utilization of computer means is presented at solving the problem from electrical engineering – solving differential equations of second order and describing time dependencies.*

Keywords: information and communication technologies, software means, MATLAB, Mathematica, MS Excel

INTRODUCTION

The most distinctive feature of the present time is implementation of information and communication technologies into everyday lives of people. These changes influence not only private (spending free time, communication) and work spheres, but also educational process. One of the most important tasks in educational area is to work out such programs and methods, at which computers would become common work tools of a teacher and at the same time they would not eliminate development of creative thinking of a student. Nowadays the issue of using computers in educational process is very often discussed at various levels. As a result of those discussions, it can be said that computers form reliable and attractive environment for learning, provide positive feedback, help to create shapely correct text, respect individual requirements , pace, speed and skills, allow to return to the problem and start or finish work in various places, help students with specific disorders of learning and handicapped students to learn, make rich information sources available, comprehensibly present complex advancement of thoughts and relations by means of graphics, offer environment for development of students thinking [1], [2].

Information and communication technologies support non-traditional forms of education (e.g. e-learning) and can contribute to development of lifelong education that is inevitable for continuous renewing and gaining necessary knowledge and skills for life in digital world. Using information and communication technologies means possibility to improve learning and thinking in many ways [3], [4].

Information technologies present one of the factors, due to which mathematical education is continuously changed, transformed and modernized. Teaching natural science and technical subjects using mathematical apparatus at technical schools is not easy and the task of pedagogues is to make educational process interesting, extraordinary and attractive.

In the following part of the paper the possibility to make teaching the subject of electric engineering by means of computer support of MS Excel, Matlab and Mathematica programs more effective is presented by particular example of solving transient performance in RLC circuit.

1. PHYSICAL ANALYSIS OF THE PROBLEM

Problem: Calculate and depict the course of voltage and current in the capacitor in RLC circuit with parameters $R = 5\Omega$, $L = 0{,}1H$, $C = 100\mu F$ connected to voltage $u = 10V$, if voltage in the capacitor at the switch was $0V$ and current flowing through the circuit at the moment of switching the circuit was zero [5].

Solution: It is a series electric circuit with *R*, *L*, *C* connected to harmonic voltage with initial conditions $i(t = 0s) = 0,\ u_C(t = 0s) = 0$ (Fig. 1).

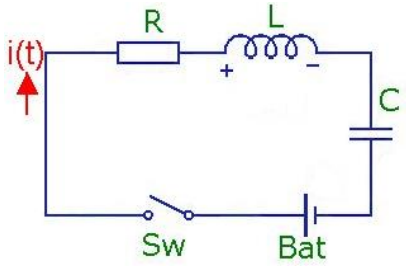


Fig. 1. RLC series circuit

Electric current, which flows through the circuit at given moment, can be determined by means of 2nd Kirchhoff´s Law. For electric voltages in the circuit we get

$$u_L + u_C + u_R = u \tag{1}$$

where u_L, u_C, u_R present momentary values of voltage on the reel with *L* induction, capacitor with *C* capacity and resistance with *R* value.
For momentary values of voltage we have

$$u_L = L\frac{di}{dt},\ \ u_C = \frac{1}{C}\int_0^t idt,\ \ u_R = Ri \tag{2}$$

and for the current flowing through capacitor we have

$$i = C\frac{du_c}{dt} \tag{3}$$

If we substitute (2), (3) equations into (1) equation we get

$$L\frac{di}{dt}+\frac{1}{C}\int_0^t idt+Ri=u$$
$$L\frac{d}{dt}\left(C\frac{du_C}{dt}\right)+\frac{1}{C}\int_0^t C\frac{du_C}{dt}dt+RC\frac{du_C}{dt}=u \quad (4)$$

We got linear integral-differential equation. For further solving it is suitable to transpose (4) equation into differential equation

$$LC\frac{d^2u_C}{dt^2}+RC\frac{du_C}{dt}+u_C=u \quad (5)$$

It is a linear differential equation of second order with constant coefficients with right side. Such a differential equation can be solved either analytically or numerically. Analytic solution of this equation requires considerable mathematical knowledge and skills. On the other side numeric solution requires computer skills (or skills with application software).

2. ANALYTICAL SOLUTION OF THE PROBLEM

General solution of the equation can be found as total of general solution of the appropriate linear differential equation with constant coefficients without right side (Y) and particular solution of linear differential equation with constant coefficients with right side (Y_P) [6] $y=Y+Y_P$.

After substitution of numeric values into (5) equation we have differential equation

$$10^{-5}\frac{d^2u_C}{dt^2}+5\cdot10^{-4}\frac{du_C}{dt}+u_C=10$$

The solution is presented in $u_C=U+U_P$. Firstly, the solution of differential equation is found. The right side of differential equation of second order with constant coefficient equals to zero and equation is solved by means of characteristic equation

$$10^{-5}\lambda^2+5\cdot10^{-4}\lambda+1=0$$

The equation is adjusted into

$$\frac{\lambda^2}{10^5}+\frac{5\lambda}{10^4}+1=0 \quad \Rightarrow \quad \lambda^2+50\lambda+10^5$$
$$\frac{\lambda^2+50\lambda+10^5}{10^5}=0$$

The solution of quadratic equation $\lambda_{1,2}=\frac{-50\pm\sqrt{(50)^2-4\cdot1\cdot10^5}}{2\cdot1}=\frac{-50\pm\sqrt{-397500}}{2}$

Since discriminant of the equation is negative, the solution is presented by two complex associated roots

$$\lambda_1=-25+25i\sqrt{159}$$
$$\lambda_2=-25-25i\sqrt{159}$$

The solution of differential equation on R is a complex function $Y=e^{\lambda x}=e^{(a+ib)x}=e^{ax}e^{ibx}=e^{ax}(\cos bx+\sin bx)$, where two functions

$y_1 = e^{ax}\cos bx, y_2 = e^{ax}\sin bx$, which are linearly independent to R and create fundamental system of equation solutions, are given. General solution is $Y = C_1 y_1 + c_2 y_2$.
General solution of the equation is as follows
$U = C_1 e^{-25t}\cos(25\sqrt{159}t) + C_2 e^{-25t}\sin(25\sqrt{159}t)$.

Particular solution (Y_P) of linear differential equation with special right side $f(x) = [R(x)\cos\beta x + S(x)\sin\beta x]e^{\alpha x}$ can be determined by the method of estimation of particular solution, where R(x) is a particular multinomial of r grade, S(x) is a particular multinomial of s grade, α, β are concrete numbers.

It is valid that $f(x) = 10 \Rightarrow \alpha = 0, \beta = 0, r = 0, s$ is not determined. For our equation we have $u_P = A$ and thus $U_P = 10$.

The entire solution of the equation can be expressed as follows
$u_c(t) = C_1 e^{-25t}\cos(25\sqrt{159}t) + C_2 e^{-25t}\sin(25\sqrt{159}t) + 10$
C_1 and C_2 constants can be determined from initial conditions $i(t = 0s) = 0,\ u_C(t = 0s) = 0$, so we have
$u_c(t = 0) = C_1 e^{-25\cdot 0}\cos(25\sqrt{159}\cdot 0) + C_2 e^{-25\cdot 0}\sin(25\sqrt{159}\cdot 0) + 10$
$u_c = C_1 + 10$
$u_c = 0 \quad \Rightarrow \quad C_1 = -10$

C_2 is determined from the following condition $i = C\frac{du_C}{dt} = 0 \quad \Rightarrow \quad \frac{du_C}{dt} = 0$.
It is valid for C_2 that

$$\frac{du_C}{dt} = C_1\left[(-25)\cdot e^{-25t}\cos\left(25\sqrt{159}t\right) + e^{-25t}\left(-\sin(25\sqrt{159}t)\right)\cdot 25\sqrt{159}\right]$$
$$+ C_2\left[(-25)e^{-25t}\sin\left(25\sqrt{159}t\right) + \left(\cos(25\sqrt{159}t)\right)\cdot 25\sqrt{159}\right] + 0$$
$$\frac{du_C}{dt}(t = 0) = C_1\left[(-25)\cdot e^{-25\cdot 0}\cos\left(25\sqrt{159}\cdot 0\right) + e^{-25\cdot 0}\left(-\sin(25\sqrt{159}\cdot 0)\right)\cdot 25\sqrt{159}\right]$$
$$+ C_2\left[(-25)e^{-25\cdot 0}\sin\left(25\sqrt{159}\cdot 0\right) + \left(\cos(25\sqrt{159}\cdot 0)\right)\cdot 25\sqrt{159}\right] + 0$$
$$\frac{du_C}{dt}(t = 0) = C_1\left[(-25) - 0\right] + C_2\left[0 + 25\sqrt{159}\right] + 0$$
$$\frac{du_C}{dt}(t = 0) = -25\cdot C_1 + 25\sqrt{159}\cdot C_2$$
$$\frac{du_C}{dt}(t = 0) = 0 \quad \Rightarrow \quad -25\cdot(-10) + 25\sqrt{159}\cdot C_2 = 0 \quad \Rightarrow \quad C_2 = -\frac{10}{\sqrt{159}}$$

Analytic solution of the differential equation (expression of dependence of voltage on time) is as follows

$$u_c(t) = -10\cdot e^{-25t}\cos(25\sqrt{159}t) - \frac{10}{\sqrt{159}}e^{-25t}\sin(25\sqrt{159}t) + 10 \qquad (6)$$

It is valid for dependence of current on time that

$$i = C\frac{du_C}{dt} = 10^{-4}\frac{d}{dt}\left(-10\cdot e^{-25t}\cos(25\sqrt{159}t) - \frac{10}{\sqrt{159}}e^{-25t}\sin(25\sqrt{159}t) + 10\right)$$

After substitution we get

$$i = \frac{4}{\sqrt{159}}e^{-25t}\sin\left(25\sqrt{159}t\right) \tag{7}$$

For dependencies of voltage and current on time the graphs of dependencies given by (6) and (7) equations are described in MS Excel program (Fig. 2 and Fig. 3).

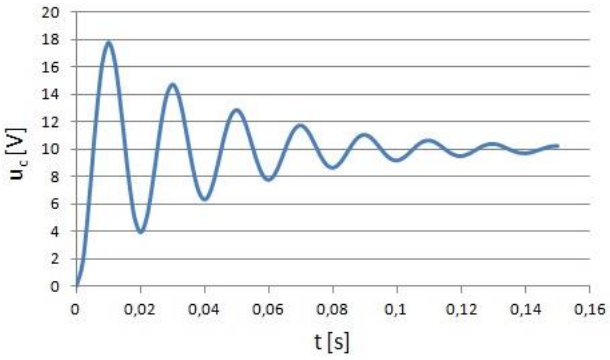


Fig. 2. Dependence of voltage on time

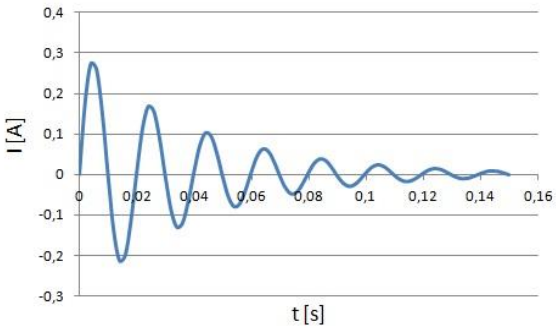


Fig. 3. Dependence of current on time

Analytic solution of the system of differential equations is difficult, it requires considerable mathematical knowledge and skills from the theory of solving differential equations and obviously does not lead to simple dependencies and results. Therefore the problem will be solved numerically by using MATLAB [7], [8].

3. NUMERICAL SOLUTION OF THE PROBLEM IN MATLAB

Efficiency of educational process can be increased by application of some modern teaching methods. One of them is implementation of information and communication technologies into teaching such as utilization of means of MATLAB at solving differential equations.

MATLAB presents highly efficient language for technical calculations. It combines calculations, visualization and programming into simply usable environment. It is an interactive tool in which the basic data type is the field without necessity to declare its parameters. This property together with number of in-built functions enables relatively easy solution of many technical problems. In school environment MATLAB is a standard tool in teaching mathematics and other technical subjects, but it is also an efficient tool for research, development and data analysis [9], [10].

MATLAB is closer to the programming language compared to other similar products. From didactic point of view it is a suitable system, because it does not require complicated programming formulae and after relatively short time a beginner can manage to work in MATLAB. On the other side it presents a strong tool for experienced users. MATLAB does not have so many prearranged mathematical functions as for example MATHEMATICA. It does not have integrated properties such as MathCad. Support of symbolic calculations is not its standard part as it is at above mentioned products. However, it does not mean that MATLAB is depleted of these possibilities. It contains more than 500 simple or more

complex mathematical functions implemented in the form of highly efficient and robust algorithms. From these functions it is possible to compose arbitrarily other functions. Sets of functions suitable for solution of a certain type of problems in MATLAB are called toolboxes. SIMULINK is an independent extension of MATLAB – solution of the system of nonlinear differential equations with a graphic entry of the system being solved. It enables graphically to observe dependencies of parameters at any connection point. It is used for simulation of dynamic behavior of the observed system. It is possible to use MATLAB in case of robust calculations, processing of extensive data files, work with large matrices and in cases when solution of the problem can be converted into vector and matrix operations. With regard to programming possibilities it is advantageous to use MATLAB also in case of branched or iterative algorithms of solution.

It is necessary to realize at the solution of differential equations of higher order in MATLAB that every differential equation of higher order can be transposed to the equivalent set of differential equations of first order with known initial conditions. At the problem solution it is suitable to transpose the differential equation of second order (5) to the set of differential equations of first order (8) as follows

$$\begin{aligned} \frac{du_C}{dt} &= \frac{i}{C} \\ \frac{di}{dt} &= \frac{u - u_C - R.i}{L} \end{aligned} \tag{8}$$

Basic standard function for the solution of differential equations is *ode45* function, which syntax is:

[*t*,*y*] = *ode*45 (*´name of the_function´, time_interval, initial_conditions*)

where *name of the_function* is reference to the function describing the set of differential equations, the parameter of *time_interval* is presented by the vector with two elements – initial time of solution t_0 and final time of solution t, the parameter of *initial_conditions* is presented by the vector of initial conditions y_0 from which we find $y(t_0) = y_0$. Two parameters are the output of the *ode45* function: t - the vector that contains instants of time, in which solution values are determined and y - the matrix containing its own solutions. To depict the current and voltage dependence on time, program writing in MATLAB is used, where the initial problem parameters, time and properties of depicted voltage and current dependences are given.

The following initial parameters are used at the problem solution: $R = 5\Omega$, $L = 10^{-1} H$, $C = 10^{-4} F$, $u = 10V$.

```
function [] = RLC_obvod
U0 = 10; % V
R = 5; % ohm
C = 1e-4; % F
L = 1e-1; % H
Uc0 = 0; % V
I0 = 0; % A
t_konec = 1.5e-1;
% duc/dt = i/C
% di/dt = (U0 - uc - Ri)/L
function dxdt = dif_rce(t, x)
    dxdt = zeros(2,1);
    dxdt(1) = x(2)/C;
    dxdt(2) = (U0 - x(1) - R*x(2))/L;
end
[t, x] = ode45(@dif_rce, [0, t_konec], [Uc0, I0]);
uc = x(:,1);
i = x(:,2);
ur = R*i;
ul = U0 - uc - ur;
subplot(1, 2, 1);
plot(t, uc, 'r');
grid on;
title('The time dependence of voltage in the circuit ');
xlabel('t [s]'); ylabel('uc [V]');
subplot(1, 2, 2);
plot(t, i, 'b');
grid on;
title('The time dependence of current in the circuit ');
xlabel('t [s]'); ylabel('i [A]');
end
```

The result of the program initialization is depiction of time dependence of the voltage and current of the RLC series circuit (Fig. 4)

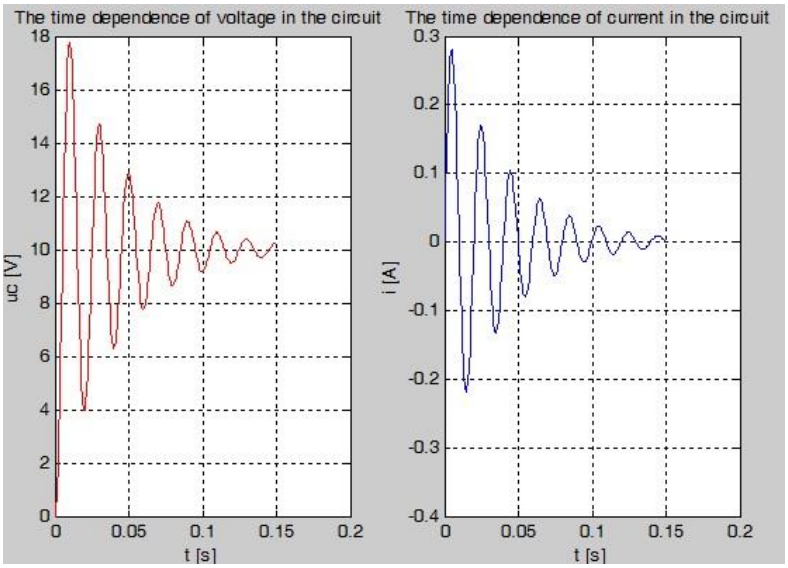


Fig. 4. Voltage and current flow in the RLC circuit

4. THE PROBLEM SOLUTION IN MATHEMATICA PROGRAM

Mathematica is a computational software program used in scientific, engineering, and mathematical fields and other areas of technical computing, which we use to solve the problem. The parameters of the problem are substituted into the equation (4) and equation is written into the Mathematica program as $(10\text{^-}5)(d(du/dt)/dt)+(5*10\text{^-}4)(du/dt)+u=10$.
After the program initialization we have the solution for voltage:

$$u_C(t)=C_1e^{-25t}\cos\left(25\sqrt{159}t\right)+C_2e^{-25t}\sin\left(25\sqrt{159}t\right)+10$$

where the solution of the equation is the sum of general solution of the appropriate equation without the right side u_c and optional particular solution u_P, t. j. $u_C(t) = u_c + u_P$. The C_1, C_2 constants are determined on the basis of the initial conditions. The result of the solution of our differential equation is (6)

$$u_C(t) = -10e^{-2500t}\cos\left(2500\sqrt{159}t\right) - \frac{10}{\sqrt{159}}e^{-2500t}\sin\left(2500\sqrt{159}t\right) + 10$$

The result of the problem solution after substitution of initial parameters in Mathematica program can be found in Fig. 5.

Solve $\frac{\frac{du(x)}{dx}}{2000} + \frac{\frac{d^2u(x)}{dx^2}}{100000} + u(x) = 10$, such that $u(0) = 0$:

The general solution will be the sum of the complementary solution and particular solution.

Find the complementary solution by solving $\frac{\frac{d^2u(x)}{dx^2}}{100000} + \frac{\frac{du(x)}{dx}}{2000} + u(x) = 0$:

Assume a solution will be proportional to $e^{\lambda x}$ for some constant λ.

Substitute $u(x) = e^{\lambda x}$ into the differential equation:

$$\frac{\frac{d^2}{dx^2}\left(e^{\lambda x}\right)}{100000} + \frac{\frac{d}{dx}\left(e^{\lambda x}\right)}{2000} + e^{\lambda x} = 0$$

Substitute $\frac{d^2}{dx^2}\left(e^{\lambda x}\right) = \lambda^2 e^{\lambda x}$ and $\frac{d}{dx}\left(e^{\lambda x}\right) = \lambda e^{\lambda x}$:

$$\frac{\lambda^2 e^{\lambda x}}{100000} + \frac{\lambda e^{\lambda x}}{2000} + e^{\lambda x} = 0$$

Factor out $e^{\lambda x}$:

$$\left(\frac{\lambda^2}{100000} + \frac{\lambda}{2000} + 1\right)e^{\lambda x} = 0$$

Since $e^{\lambda x} \neq 0$ for any finite λ, the zeros must come from the polynomial:

$$\frac{\lambda^2}{100000} + \frac{\lambda}{2000} + 1 = 0$$

Factor:

$$\frac{\lambda^2 + 50\lambda + 100000}{100000} = 0$$

Solve for λ:

$\lambda = -25 + (25\,i)\sqrt{159}$ or $\lambda = -25 - (25\,i)\sqrt{159}$

The roots $\lambda = -25 \pm 25\,i\sqrt{159}$ give $u_1(x) = c_1 e^{\left(-25+25\,i\sqrt{159}\right)x}$, $u_2(x) = c_2 e^{\left(-25-25\,i\sqrt{159}\right)x}$ as solutions, where c_1 and c_2 are arbitrary constants.

The general solution is the sum of the above solutions:

$$u(x) = u_1(x) + u_2(x) = c_1 e^{\left(-25+(25\,i)\sqrt{159}\right)x} + c_2 e^{\left(-25-(25\,i)\sqrt{159}\right)x}$$

Apply Euler's identity $e^{\alpha+i\beta} = e^{\alpha}\cos(\beta) + i\,e^{\alpha}\sin(\beta)$:

$$u(x) = c_1\left(\frac{\cos\left(25\sqrt{159}\,x\right)}{e^{25x}} + \frac{i\sin\left(25\sqrt{159}\,x\right)}{e^{25x}}\right) + c_2\left(\frac{\cos\left(25\sqrt{159}\,x\right)}{e^{25x}} - \frac{i\sin\left(25\sqrt{159}\,x\right)}{e^{25x}}\right)$$

Regroup terms:

$$u(x) = \frac{(c_1 + c_2)\cos\left(25\sqrt{159}\,x\right)}{e^{25x}} + \frac{i\,(c_1 - c_2)\sin\left(25\sqrt{159}\,x\right)}{e^{25x}}$$

Redefine $c_1 + c_2$ as c_1 and $i\,(c_1 - c_2)$ as c_2, since these are arbitrary constants:

$$u(x) = \frac{c_1\cos\left(25\sqrt{159}\,x\right)}{e^{25x}} + \frac{c_2\sin\left(25\sqrt{159}\,x\right)}{e^{25x}}$$

Determine the particular solution to $\frac{\frac{d^2u(x)}{dx^2}}{100000} + u(x) + \frac{\frac{du(x)}{dx}}{2000} = 10$ by the method of undetermined coefficients:

The particular solution to $\frac{\frac{d^2u(x)}{dx^2}}{100000} + u(x) + \frac{\frac{du(x)}{dx}}{2000} = 10$ is of the form:

$u_p(x) = a_1$

Solve for the unknown constant a_1:

Compute $\frac{du_p(x)}{dx}$:

$$\frac{du_p(x)}{dx} = \frac{d}{dx}(a_1) = 0$$

Compute $\frac{d^2u_p(x)}{dx^2}$:

$$\frac{d^2u_p(x)}{dx^2} = \frac{d^2}{dx^2}(a_1) = 0$$

Substitute the particular solution $u_p(x)$ into the differential equation:

$$\frac{\frac{d^2u_p(x)}{dx^2}}{100000} + \frac{\frac{du_p(x)}{dx}}{2000} + u_p(x) = 10$$

$$\frac{0}{100000} + \frac{0}{2000} + a_1 = 10$$

Substitute a_1 into $u_p(x) = a_1$:

$u_p(x) = 10$

The general solution is:

$$u(x) = u_c(x) + u_p(x) = \frac{c_1\cos\left(25\sqrt{159}\,x\right)}{e^{25x}} + \frac{c_2\sin\left(25\sqrt{159}\,x\right)}{e^{25x}} + 10$$

Solve for the unknown constants using the initial conditions:

Substitute $u(0) = 0$ into $u(x) = c_2 e^{-25x}\sin\left(25\sqrt{159}\,x\right) + c_1 e^{-25x}\cos\left(25\sqrt{159}\,x\right) + 10$:

$c_1 + 10 = 0$

Solve the equation:

$c_1 = -10$

Substitute $c_1 = -10$ into $u(x) = c_2 e^{-25x}\sin\left(25\sqrt{159}\,x\right) + c_1 e^{-25x}\cos\left(25\sqrt{159}\,x\right) + 10$:

Answer:

$$u(x) = -\frac{10\cos\left(25\sqrt{159}\,x\right)}{e^{25x}} + \frac{c_2\sin\left(25\sqrt{159}\,x\right)}{e^{25x}} + 10$$

Fig. 5. Solution of differential equation in Mathematica program

The graph of voltage dependence can be found in Fig. 6.

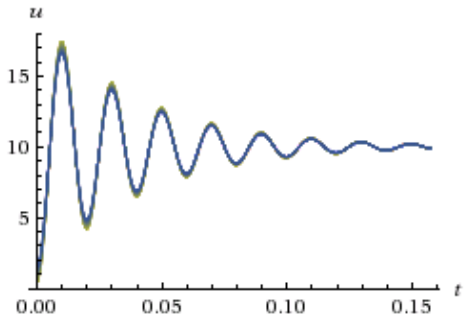


Fig. 6. Voltage characteristics of RLC circuit

For the current we have:

$$i(t) = C\frac{du_C}{dt} = \frac{400e^{-2500t}\sin\left(2500\sqrt{159}t\right)}{\sqrt{159}} \tag{8}$$

The equation (8) is written into the Mathematica program as follows $400*(e\text{^}\text{-}(2500*t))*\sin(2500*sqrt(159)*t)/sqrt(159) = i$. The graph of current dependence can be found in Fig. 7.

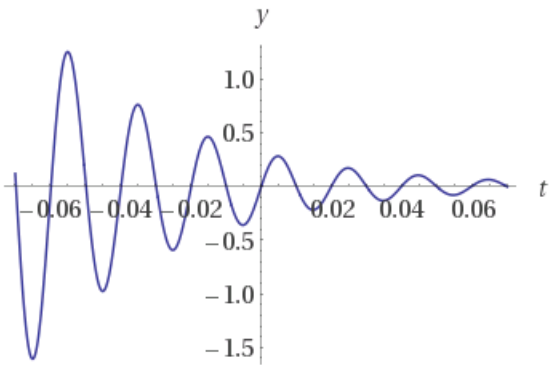


Fig. 7. Current dependence of RLC circuit

CONCLUSION

The presence of information and communication technologies in teaching has positive influence on efficiency of educational process and students accept them very positively. Based on the results it can be said that it is possible to facilitate and improve educational process by suitable combination of classic teaching methods and introduction of new elements using information and communication means. Gradual introduction and utilization of computer technique into teaching natural science and technical subjects at universities is one of the significant elements of modernizing technical and natural science education. The main goal of these modernizing tendencies is continual increase of educational level corresponding to present-day degree of knowledge in individual teaching disciplines with the respect to the age of a student and type of school studied. On the other side teaching these subjects with the support of computer technique cannot be understood in any case as a universal way how to solve the problems in educational system.

REFERENCES

[1] KALAŠ, I. *Čo ponúkajú informačné a komunikačné technológie iným predmetom*. Bratislava: ŠPÚ, 2001.

[2] TUREK, I. *Zvyšovanie efektívnosti vyučovania*. MC Bratislava, 1997.

[3] KIREŠ, Marián. Informačno-komunikačné technológie a potreba ďalšieho vzdelávania učiteľov. In: *Medzinárodná vedecko – odborná konferencia „Fyzika v kontexte kultúry“*. Prešov: FHPV Prešovská univerzita, 2001, pp. 235-238. ISBN 80-8068-082-5.

[4] BRESTENSKÁ, B. a kol. *Premena školy s využitím informačných a komunikačných technológií.* Elfa, s. r. o. Košice, 2010. ISBN 978-80-8086-143-8.

[5] HAJKO, Vladimír, DANIEL-SZABÓ Juraj. Základy fyziky. Bratislava: Veda, 1983.

[6] VAGASKÁ, Alena, MIŽÁKOVÁ, Jana. MATEMATIKA II. TU v Košiciach, Prešov, 2014. ISBN: 978-80-553-1670-3.

[7] HREHOVÁ, Stella. Approximations of the solution of the differential equation using the choosen application programs. In: *20. DIDMATTECH 2007*. Olomouc: Votobia, 2007, pp. 323-326. ISBN 80-7220-296-0.

[8] VAGASKÁ, Alena. Computer support of applied mathematics in the engineering study programe. In: *6. konference o matematice a fyzice na vysokých školách technických*. Brno: Univerzita obrany, 2009, pp. 309-316. ISBN 978-80-7231-668-7.

[9] KARBAN, Pavel. Výpočty a simulace v programoch Matlab a Simulink. Brno: Computer Press, a.s, 2007, 224 pp. ISBN 80-2511-448-1.

[10] DUŠEK, František. MATLAB a SIMULINK- úvod do používání. Pardubice: Univerzita Pardubice, 2002, 158 pp. ISBN 80-7194-273-1.

Acknowledgement
The work presented in this paper has been supported by the project VEGA, No. 1/0738/14.

CONSTRUCTING SOLUTIONS OF LINEAR STATIONARY EQUATION OF THE SECOND ORDER DELAY

Khusainov D.Ya., Dzhalladova I.A., Pokoyovy M.V.
Kyiv Taras Shevchenko National University of Kyiv, Kyiv Vadim Getman National Ekonomical University of Kyiv, University of Konstanz
d.y.khusainov@gmail.com, idzhalladova@gmail.com, michael.pokoyovy@uni-konstanz.de

Abstract: *In this work we review the construction of solutions of linear equations of second order stationary with delay. We have special functions called lagging exponentials, and there combinations by which obtained a solution of the Cauchy problem. We studied real eigenvalues of different signs, eigenvalues of complex conjugate eigenvalues, eigenvalues of really different.*

Keywords: differential equation, delay, solutions

INTRODUCTION

Let's consider the linear differential equation of second order with constant coefficients. In the absence of delay, it has the form

$$x''(t)+ px'(t)+ qx(t)= f(t)$$

and finding solutions of the Cauchy problem $x(0)= x_0,\ x'(0)= x_0'$, the equation reduces to the investigation of the roots of the characteristic equation

$$\lambda^2 + p\lambda + q = 0.$$

In this report we reviews the differential equations with delay

$$x''(t)+ px'(t-\tau)+ qx(t-2\tau)= f(t).$$

The characteristic equation corresponding to the equation has the form $\lambda^2 + p\lambda e^{-\lambda\tau} + qe^{-2\lambda\tau} = 0.$

This transcendental equation, and it has a countable number of roots.

We have special functions called lagging exponentials, and there combinations by which obtained a solution of the Cauchy problem $x(t)=\varphi(t),\ x'(t)=\varphi'(t),\ -2\tau \le t \le 0$. .

1. THE EIGENVALUES ARE REAL, DIFFERENT SIGNS.

We consider the equation without delay

$$x''(t)-\Omega^2 x(t)= f(t),\ t \ge 0. \tag{1.1}$$

The solution of the homogeneous equation which satisfying the initial conditions

$$x(0)= x_0,\ x'(0)= x_0'. \tag{1.2}$$

It has the form

$$x_{od}(t)=\frac{1}{2}x_0\left(e^{\Omega t}+e^{-\Omega t}\right)+\frac{1}{2\Omega}\left(e^{\Omega t}-e^{-\Omega t}\right). \tag{1.3}$$

A particular solution of the inhomogeneous equation which satisfying zero initial conditions, we search in the form of Cauchy

$$x_{ch}(t)=\int_0^t K(t,s)f(s)ds, \tag{1.4}$$

Cauchy kernel has the form $K(t,s)=\frac{1}{2\Omega}\left(e^{\Omega(t-s)}-e^{-\Omega(t-s)}\right)$,

And the solution of the Cauchy problem of the inhomogeneous equation (1.1) with non-zero initial conditions (1.2) has the form

$$x(t)=\frac{1}{2}x_0\left(e^{\Omega t}+e^{-\Omega t}\right)+\frac{1}{2\Omega}x_0'\left(e^{\Omega t}-e^{-\Omega t}\right)+\frac{1}{2\Omega}\int_0^t\left(e^{\Omega(t-s)}-e^{-\Omega(t-s)}\right)f(s)ds. \qquad (1.5)$$

If we denote $x_1(t)=\frac{1}{2}\left(e^{\Omega t}+e^{-\Omega t}\right)$, $x_2(t)=\frac{1}{2\Omega}\left(e^{\Omega t}-e^{-\Omega t}\right)$,

the dependence takes the form

$$x(t)=x_0x_1(t)+x_0'x_2(t)+\int_0^t x_2(t-s)f(s)ds. \qquad (1.6)$$

The equation of the second order with delay. We consider the equation with one constant delay

$$x''(t)-\Omega^2x(t-2\tau)=f(t),\ t>0 \qquad (1.7)$$

We get it solution that satisfies the initial conditions

$$x(t)=\varphi(t),\ x'(t)=\varphi'(t),\ -2\tau\le t\le 0. \qquad (1.8)$$

It is shown that there is a representation of the solution in a form similar to the dependence (1.6).

Definition 1.1. A lagging exponential $\exp_\tau\{\Omega,t\}$ with indicator Ω and delay τ is a function that has the form

$$\exp_\tau\{\Omega,t\}=\begin{cases}0, & -\infty<t<-\tau,\\ 1, & -\tau\le t<0,\\ 1+\Omega\frac{t}{1!}, & 0\le t<\tau,\\ 1+\Omega\frac{t}{1!}+\Omega^2\frac{(t-\tau)^2}{2!}, & \tau\le t<2\tau\\ \ldots, & \ldots\\ 1+\Omega\frac{t}{1!}+\cdots+\Omega^k\frac{[t-(k-1)\tau]^k}{k!}, & (k-1)\tau\le t<k\tau\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \end{cases} \qquad (1.9)$$

We introduce two functions which are linear combinations of exponentials delayed.

$$x_1(t)=\frac{1}{2}\left[\exp_\tau\{\Omega,t\}+\exp_\tau(-\Omega,t)\right],\quad x_2(t)=\frac{1}{2\Omega}\left[\exp_\tau\{\Omega,t\}-\exp_\tau(-\Omega,t)\right].$$

As the presentation of the delayed exponential $\exp_\tau\{\Omega,t\}$, we have the following dependence

$$x_1(t)=\begin{cases}0, & -\infty<t<-\tau,\\ 1, & -\tau\le t<\tau,\\ 1+\Omega^2\frac{(t-\tau)^2}{2!}, & \tau\le t<3\tau,\\ 1+\Omega^2\frac{(t-\tau)^2}{2!}+\Omega^4\frac{(t-3\tau)^4}{4!}, & 3\tau\le t<5\tau\\ \ldots, & \ldots\\ 1+\Omega^2\frac{(t-\tau)^2}{2!}+\cdots+\Omega^{2k}\frac{[t-(2k-1)\tau]^{2k}}{(2k)!}, & (2k-1)\tau\le t<(2k+1)\tau\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \end{cases} \qquad (1.10)$$

$$x_2(t)=\begin{cases} 0 & , & -\infty < t < 0, \\ \frac{t}{1!} & , & 0 \le t < 2\tau, \\ \frac{t}{1!}+\Omega^2\frac{(t-2\tau)^3}{3!} & , & 2\tau \le t < 4\tau, \\ \frac{t}{1!}+\Omega^2\frac{(t-2\tau)^3}{3!}+\Omega^4\frac{(t-4\tau)^5}{5!} & , & 4\tau \le t < 6\tau \\ \cdots & , & \cdots \\ \frac{t}{1!}+\Omega^2\frac{(t-2\tau)^3}{3!}+\cdots+\Omega^{2k}\frac{[t-(2k)\tau]^{2k+1}}{(2k+1)!} & , & (2k)\tau \le t < (2k+2)\tau \\ \cdots\cdots\cdots & & \end{cases} \qquad (1.11)$$

We received the following statement.

Theorem 1.1. The solution of the Cauchy problem (1.2) for the homogeneous equation with delay can be written as

$$x(t)=\varphi(-2\tau)x_1(t+\tau)+\varphi'(-2\tau)x_2(t+2\tau)+\int_{-2\tau}^{0}x_2(t-s)\varphi''(s)ds, \qquad (1.12)$$

Where $x_1(t)$ represented in (1.4) and $x_2(t)$ presented in (1.5).

Theorem 1.2. The solution of the Cauchy problem with zero initial conditions $x(t)=0$, $x'(t)=0$, $-2\tau \le t \le 0$ for the inhomogeneous equation has the form

$$x(t)=\int_0^t x_2(t-s)f(s)ds,\ t \ge \tau, \qquad (1.13)$$

2. THE EIGENVALUES OF THE COMPLEX CONJUGATE.

Let the eigenvalues of the characteristic equation are equal

$$\lambda_1=\frac{1}{2}\left(-p_1+\sqrt{p_1^2-4p_2}\right),\ \lambda_2=\frac{1}{2}\left(-p_1-\sqrt{p_1^2-4p_2}\right),$$

$p_1^2 < 4p_2$ and eigenvalues of the complex conjugate, i.e. $\lambda_{1,2}=p \pm iq$, $p=-\frac{1}{2}p_1$, $q=\frac{1}{2}\sqrt{4p_2-p_1^2}$. Then the general solution of the homogeneous equation is given without delay

$$x_{od}(t)=x_1(t)x_0+x_2(t)x_0',\ x_1(t)=e^{pt}\left(\cos qt-\frac{p}{q}\sin qt\right),\ x_2(t)=\frac{1}{q}e^{pt}\sin qt. \qquad (2.1)$$

a particular solution satisfying zero initial conditions $x_{ch}(t)=\int_0^t x_2(t-s)f(s)ds$.

Definitively, the solution of the Cauchy problem of the inhomogeneous equation (1.1) with non-zero initial conditions (1.2) has the form

$$x(t)=x_1(t)x_0+x_2(t)x_0'+\int_0^t x_2(t-s)f(s)ds. \qquad (2.2)$$

The equation of the second order with delay. Consider the homogeneous differential equation with delay

$$x''(t)+p_1x'(t-\tau)+p_2x(t-2\tau)=0,\ t \ge 0. \qquad (2.3)$$

Definition 2.1. A lagging exponential $\exp_\tau\{\lambda_1,t\}$ with indicator of the complex $\lambda_1=p+iq$ and delay τ will be the function that has the form

$$\exp_\tau\{\lambda_1,t\}=\begin{cases}0 & , \quad -\infty<t<-\tau,\\ 1 & , \quad -\tau\le t<0,\\ 1+(p+iq)\frac{t}{1!} & , \quad 0\le t<\tau,\\ 1+(p+iq)\frac{t}{1!}+(p+iq)^2\frac{(t-\tau)^2}{2!} & , \quad \tau\le t<2\tau\\ \cdots & , \quad \cdots\\ 1+(p+iq)\frac{t}{1!}+\cdots+(p+iq)^k\frac{[t-(k-1)\tau]^k}{k!} & , \quad (k-1)\tau\le t<k\tau\\ \cdots\cdots\cdots & \end{cases} \tag{2.4}$$

Let's consider auxiliary statement.

Lemma 2.1. occurs of ratio

$$(p+iq)^k=r^k\cos k\varphi+ir^k\sin k\varphi,\ r=\sqrt{p^2+q^2},\ \varphi=arctg\frac{q}{p}. \tag{2.5}$$

Lemma 2.2. Delayed exponential $\exp_\tau\{\lambda_1,t\}$ (with the index $\lambda_1=p+iq$ and delay τ) can be written in the form of a complex function $\exp_\tau\{\lambda_1,t\}=u_\tau\{p,q,t\}+iv_\tau\{p,q,t\}$,

Where

$$u_\tau\{p,q,t\}=\begin{cases}0 & , \quad -\infty<t<-\tau,\\ 1 & , \quad -\tau\le t<0,\\ 1+r\cos\varphi\frac{t}{1!} & , \quad 0\le t<\tau,\\ 1+r\cos\varphi\frac{t}{1!}+r^2\cos 2\varphi\frac{(t-\tau)^2}{2!} & , \quad \tau\le t<2\tau\\ \cdots & , \quad \cdots\\ 1+r\cos\varphi\frac{t}{1!}+\cdots+r^k\cos k\varphi\frac{[t-(k-1)\tau]^k}{k!} & , \quad (k-1)\tau\le t<k\tau\\ \cdots\cdots\cdots & \end{cases} \tag{2.6}$$

$$v_\tau\{p,q,t\}=\begin{cases}0 & , \quad -\infty<t<-\tau,\\ 0 & , \quad -\tau\le t<0,\\ r\sin\varphi\frac{t}{1!} & , \quad 0\le t<\tau,\\ r\sin\varphi\frac{t}{1!}+r^2\sin 2\varphi\frac{(t-\tau)^2}{2!} & , \quad \tau\le t<2\tau\\ \cdots & , \quad \cdots\\ r\sin\varphi\frac{t}{1!}+\cdots+r^k\sin k\varphi\frac{[t-(k-1)\tau]^k}{k!} & , \quad (k-1)\tau\le t<k\tau\\ \cdots\cdots\cdots & \end{cases} \tag{2.7}$$

Lemma 2.3. Delayed exponential $\exp_\tau\{\lambda_2,t\}$, (with the index $\lambda_2=p-iq$ and delay τ) can be written in the form of a complex function $\exp_\tau\{\lambda_2,t\}=u_\tau\{p,q,t\}-iv_\tau\{p,q,t\}$.

Theorem 2.1. Delayed exponential $\exp_\tau\{\lambda_1,t\}$ with the index

$$\lambda_1=p+iq,\ p=-\frac{1}{2}p_1,\ q=\frac{1}{2}\sqrt{4p_2-p_1^2} \tag{2.9}$$

it is the solution of differential equationa with delay (2.1) satisfies the initial condition

$$x(t)\equiv 1,\ -\tau\le t\le 0. \tag{2.10}$$

Similarly, we can prove.

Theorem 2.2. Delayed exponential $\exp_\tau\{\lambda_2,t\}$ with an exponent

$$\lambda_2=p-iq,\ p=-\frac{1}{2}p_1,\ q=\frac{1}{2}\sqrt{4p_2-p_1^2} \tag{2.11}$$

It is the solution of differential equation with delay(2.1) which satisfying the initial conditions (2.10).

We introduce two functions wich is a linear combination of delayed exponentials $\exp_\tau\{\lambda_1, t\}$, $\exp_\tau\{\lambda_2, t\}$.

$$x_1(t) = \frac{1}{\lambda_2 - \lambda_1}\left[\lambda_2 \exp_\tau\{\lambda_1, t\} - \lambda_1 \exp_\tau\{\lambda_2, t\}\right], \; x_2(t) = \frac{1}{\lambda_2 - \lambda_1}\left[\exp_\tau\{\lambda_2, t\} - \exp_\tau(\lambda_1, t)\right]. \quad (2.12)$$

Here the parameters λ_1, λ_2 are defined in (2.9), (2.11).
As follows from the representation (2.3) of the delayed exponential $\exp_\tau\{\lambda, t\}$, the following dependence takes place

$$x_1(t) = \begin{cases} 0 & , \quad -\infty < t < -\tau, \\ 1 & , \quad -\tau \le t < 0, \\ 1 & , \quad 0 \le t < \tau, \\ 1 + \lambda_1\lambda_2 \dfrac{\lambda_1^2 - \lambda_2^2}{\lambda_2 - \lambda_1} \times \dfrac{(t-\tau)^2}{2!} & , \quad \tau \le t < 2\tau, \\ \cdots & , \quad \cdots \\ 1 + \lambda_1\lambda_2 \dfrac{\lambda_1 - \lambda_2}{\lambda_2 - \lambda_1}\dfrac{(t-\tau)^2}{2!} + \cdots + \lambda_1\lambda_2 \dfrac{\lambda_1^k - \lambda_2^{k1}}{\lambda_2 - \lambda_1}\dfrac{[t-(k-1)\tau]^k}{(k)!} & , \quad (k-1)\tau \le t < k\tau, \\ \cdots\cdots\cdots & \end{cases} \quad (2.11)$$

forasmuch as $\lambda_1\lambda_2 = r^2$, $\lambda_1^k - \lambda_2^k = 2ir^k \sin k\varphi$, $k = 1,2,3,\ldots$, then

$$\lambda_1\lambda_2 \frac{\lambda_1^{k-1} - \lambda_2^{k-1}}{\lambda_2 - \lambda_1} = -r^k \frac{\sin(k-1)\varphi}{\sin\varphi}.$$

Here we obtain

$$x_1(t) = \begin{cases} 0 & , \quad -\infty < t < -\tau, \\ 1 & , \quad -\tau \le t < 0, \\ 1 & , \quad 0 \le t < \tau, \\ 1 - r^2 \dfrac{(t-\tau)^2}{2!} & , \quad \tau \le t < 2\tau, \\ \cdots & , \quad \cdots \\ 1 + r^2 \dfrac{(t-\tau)^2}{2!} + \cdots - r^k \dfrac{\sin(k-1)\varphi}{\sin\varphi} \times \dfrac{[t-(k-1)\tau]^k}{(k)!} & , \quad (k-1)\tau \le t < k\tau, \\ \cdots\cdots\cdots & \end{cases}$$

Similarly for the function $x_2(t)$

$$x_2(t) = \begin{cases} 0 & , \quad -\infty < t < -\tau, \\ 0 & , \quad -\tau \le t < 0, \\ \dfrac{t}{1!} & , \quad 0 \le t < \tau, \\ \dfrac{t}{1!} + r\dfrac{\sin 2\varphi}{\sin\varphi} \times \dfrac{(t-\tau)^2}{2!} & , \quad \tau \le t < 2\tau \\ \dfrac{t}{1!} + r\dfrac{\sin 2\varphi}{\sin\varphi} \times \dfrac{(t-\tau)^2}{2!} + \dfrac{\sin 3\varphi}{\sin\varphi} \times \dfrac{(t-2\tau)^3}{3!} & , \quad 2\tau \le t < 3\tau \\ \cdots & , \quad \cdots \\ \dfrac{t}{1!} + \dfrac{\sin 2\varphi}{\sin\varphi} \times \dfrac{(t-\tau)^2}{2!} + \cdots + \dfrac{\sin k\varphi}{\sin\varphi} \times \dfrac{[t-(k-1)\tau]^k}{k!} & , \quad (k-1)\tau \le t < k\tau \\ \cdots\cdots\cdots & \end{cases} \quad (2.12)$$

Corollary 2.1. A linear combination $x_1(t)$ with delayed exponentials is the solution of the differential equation with delay (2.1).

Corollary 2.2. A linear combination $x_2(t)$ of delayed exponentials is a solution of the differential equation with delay (2.3) wich satisfying the initial condition

$$x(t) \equiv t,\ 0 \le t \le \tau. \tag{2.13}$$

Based on the above allegations we get Representation of solution of the Cauchy problem for equations with complex conjugate eigenvalues.

3. THE REAL EIGENVALUES, DIFFERENT.

Let the eigenvalues of the characteristic equation

$$\lambda_1 = \frac{1}{2}\left(-p_1 + \sqrt{p_1^2 - 4p_2}\right),\ \lambda_2 = \frac{1}{2}\left(-p_1 - \sqrt{p_1^2 - 4p_2}\right).$$

Indeed, various i.e. $\lambda_1 \ne \lambda_2$. Then the general solution of the homogeneous equation without delay looks like

$$x_{od}(t) = x_1(t)x_0 + x_2(t)x_0',\ x_1(t) = \frac{\lambda_2 e^{\lambda_1 t} - \lambda_1 e^{\lambda_2 t}}{\lambda_2 - \lambda_1},\ x_2(t) = \frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1}. \tag{3.1}$$

Cauchy kernel has the form $K(t,s) = \frac{e^{\lambda_2 (t-s)} - e^{\lambda_1 (t-s)}}{\lambda_2 - \lambda_1} = x_2(t-s)$,

and the solution of the Cauchy problem of the inhomogeneous equation (1.1) with non-zero initial conditions (1.2) can be written in the same integral form.

Let's consider homogeneous differential equation with delay

$$x''(t) + p_1 x'(t-\tau) + p_2 x(t-2\tau) = 0,\ p_1 = -(\lambda_1 + \lambda_2),\ p_2 = \lambda_1\lambda_2, \lambda_1 \ne \lambda_2,\ t \ge 0. \tag{3.2}$$

with the initial conditions

$$x(t) = \varphi(t),\ x'(t) = \varphi'(t),\ -2\tau \le t \le 0. \tag{3.3}$$

We introduce two functions it is a linear combination of delayed exponentials $\exp_\tau\{\lambda_1, t\}$, $\exp_\tau\{\lambda_2, t\}$.

$$x_1(t) = \frac{1}{\lambda_2 - \lambda_1}\left[\lambda_2 \exp_\tau\{\lambda_1, t\} - \lambda_1 \exp_\tau(\lambda_2, t)\right],\ x_2(t) = \frac{1}{\lambda_2 - \lambda_1}\left[\exp_\tau\{\lambda_2, t\} - \exp_\tau(\lambda_1, t)\right]. \tag{3.4}$$

As follows from the representation (2.3) of the delayed exponential $\exp_\tau\{\lambda, t\}$, have the following depending

$$x_1(t) = \begin{cases} 0, & -\infty < t < -\tau, \\ 1, & -\tau \le t < 0, \\ 1, & 0 \le t < \tau, \\ 1 - \lambda_1\lambda_2 \frac{(t-\tau)^2}{2!} - \lambda_1\lambda_2(\lambda_1 + \lambda_2)\frac{(t-2\tau)^3}{3!}, & \tau \le t < 2\tau, \\ \cdots, & \cdots \\ 1 - \lambda_1\lambda_2 \frac{(t-\tau)^2}{2!} + \cdots + \frac{\lambda_2\lambda_1^k - \lambda_1\lambda_2^k}{\lambda_2 - \lambda_1}\frac{[t-(k-1)\tau]^k}{(k)!}, & (k-1)\tau \le t < k\tau, \\ \cdots\cdots\cdots & \end{cases} \tag{3.5}$$

$$x_2(t)=\begin{cases} 0 & , \quad -\infty < t < -\tau, \\ 0 & , \quad -\tau \le t < 0, \\ \dfrac{t}{1!} & , \quad 0 \le t < \tau, \\ \dfrac{t}{1!}+(\lambda_1+\lambda_2)\dfrac{(t-\tau)^2}{2!} & , \quad \tau \le t < 2\tau \\ \dfrac{t}{1!}+(\lambda_1+\lambda_2)\dfrac{(t-\tau)^2}{2!}+(\lambda_1^2+\lambda_1\lambda_2+\lambda_2^2)\dfrac{(t-2\tau)^3}{3!} & , \quad 2\tau \le t < 3\tau \\ \cdots & , \quad \cdots \\ \dfrac{t}{1!}+(\lambda_1+\lambda_2)\dfrac{(t-\tau)^2}{2!}+\cdots+\sum_{i=1}^{k-1}\lambda_1^{k-1-i}\lambda_2^{i}\dfrac{[t-(k-1)\tau]^k}{k!} & , \quad (k-1)\tau \le t < k\tau \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \end{cases} \qquad (3.6)$$

Corollary 3.1. A linear combination $x_1(t)$ of retarded exponentials is a solution of the homogeneous differential equation with delay satisfying the initial conditions (3.4).

Corollary 3.2. A linear combination $x_2(t)$ of delayed exponentials is a solution of the differential equation with delay satisfying the initial conditions $x(t) \equiv t$, $0 \le t \le \tau$.

With using of the above functions we obtain a solution of the Cauchy problem for the equation with delay.

REFERENCES

[1] ÈL'SGOL'TS L.E., NORKIN S.B. *Introduction to the theory of differential equations with deviating argument*. M., Nauka, 1971.

[2] J. HALE. *The theory of functional differential equations*. M., Mir, 1984. 484 pp.

[3] AZIZBAYOV E.I., KHUSAINOV D.Ya. *The solution of an equation with delay*. Bulletin of Kyiv National Taras Shevchenko University. Series: Science, B.12, 2012. pp. 4-11.

[4] KHUSAINOV D.Ya., SHUKLIN G.V. *The relative controllability in pure delay*. Journal of Applied Mechanics. 2005. 41, №2. pp.118-130.

[5] D.Khusainov, M.Pokojovy, R.Racke. *Strong and Mild Extrapolated Solutions to the Heat Equation with Constant Delay*. Konstancer Schriften in Mathematik, Konstant University, No.320. 2013. 32 pp.

THERMALLY ACTIVATED DEFORMATION AND DYNAMIC STRAIN AGING OF Cd-Zn SINGLE CRYSTALS ALLOYS

Vladislav Navrátil
Department of Physics, Chemistry and Vocational Education, Faculty of Education,
Masaryk University
Poříčí 7, 603 00 Brno, Czech Republic
navratil@ped.muni.cz

Abstract. *A piece of metal can be deformed permanently if it is pulled sufficiently hard in tension, compression or is twisted through a large enough angle in torsion. When the stress is removed, the dimensions of the piece of metal do not return to their original values as they would do if the deformation were elastic. The permanent distortion suffered by the metal specimen is called plastic deformation. The chief mechanism by which plastic deformation occurs is the motion of dislocations. Because there are an immense number of ways in which dislocations can bring about plastic deformation it is not surprising that this phenomenon is quite complex. The character of plastic deformation is a sensitive function of such variables as temperature, the strain rate of deformation, the past history of the sample, crystal size, and if the sample is a single crystal, the orientation of the axes with respect to the stress system. Much interest has recently been taken also in the influence which small quantities of foreign elements may have on the properties of metals. In fact, impurities play an important part in metal physics research. They form a particular species of point defects, and are able to interact with the other lattice defects which exist in the metal and determine a great number of its properties.*

It is the purpose of the present work to examine the temperature and solute atoms concentration influence on mechanical properties of Cd-Zn single crystals alloys. *The author wish to express his thanks to Prof. Dr. Pavel Lukáč, DrSc., and Doc.Dr.Miloš Hamerský, CSc For the valuable discussions which have preceded the original of this work.*

Keywords:, thermodynamics, dislocations, creep, stress exponent, dynamic strain aging.

1. INTRODUCTION

There exist many ways of producing the plastic deformation of solids. One of the simplest and most applicable is the deformation by a tensile force, the so-called tensile test. In the present paper we shall investigate the plastic behavior of Cd-Zn single crystals by means of then tensile test, called creep. Creep is a tensile test where a specimen undergoes a continuous deformation under a constant load or stress.

When a solid is subjected to a static force, the atomic lattice will adjust itself to oppose the applied force and maintain equilibrium. On a macroscopic scale the atomic adjustment is observed as a deformation which can be measured macroscopically. The deformation referred to unit elements of the length of the sample and thus converted into the dimensionless quantity is called "strain ε". The response of strain to the applied stress σ varies with the

magnitude of this stress, temperature and strain rate. Experimental evidence conclusively shows that the creep flow is thermally activated. It means that the local thermal agitation provides additional energy, beyond that provided mechanically, to overcome barriers to creep deformation. From the physical viewpoint it means that creep is a suitable method for investigating these processes, because the plastic deformation can occur under a constant stress owing to these processes only. The applied stress aids in overcoming these barriers and serves to give direction to the resultant flow.

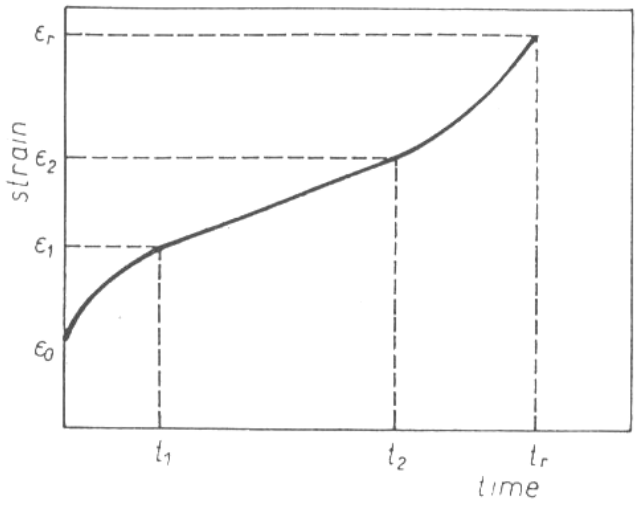
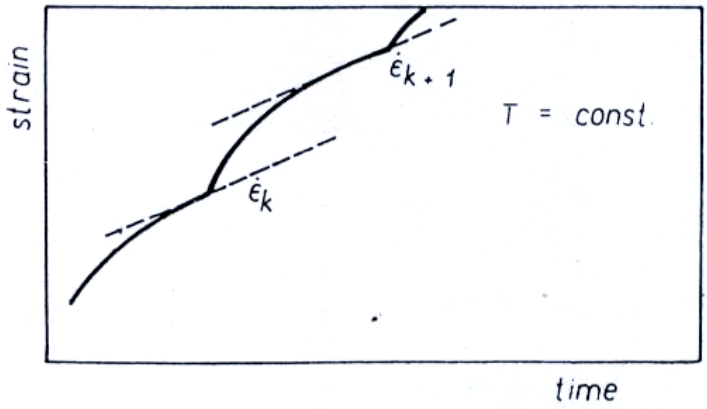


Fig.1. Schematic representation of a creep- rupture curves

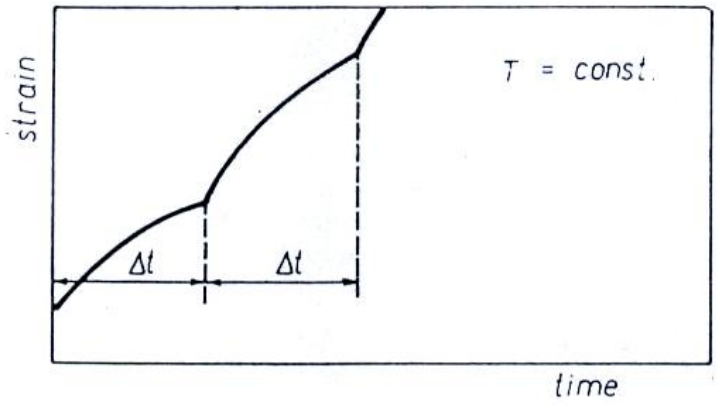
The creep of metals can be demonstrated directly by a creep curve which represents graphically the function between creep strain and time. An idealized creep curve is shown schematically in Fig.1. The strain ε_0 is obtained immediately upon loading and exhibits characteristics of plastic deformation, but, of course, also includes elastic deformation. Between ε_0 and ε_1 the creep rate decreases continually. This period of the creep curve is called "primary creep". Between ε_1 and ε_2 the creep rate remains nearly constant, indicating a nearly steady-state condition. This part of the creep history in which the strain rate $d\varepsilon/dt$ remains nearly constant is called "secondary creep" or "steady state creep". Beyond ε_2 the creep rate increases until rupture occurs at the strain ε_r and rupture time t_r. The period of increasing creep rate is called "tertiary creep". In the present paper we shall mostly deal with the primary (transient) creep of Cd-Zn alloys. The suitable method for studying the transient creep it is the incremental loading method, which has been used by a number of authors [1-5, 10-12]. There are two variants of this method:

a) The creep deformation occurs by incremental loading with increments so that the next increment is added when the strain rate decreases to value fixed before (Fig.2a).
b) The sample is gradually loaded with increments in constant time intervals (Fig.2b).

We used the second one method and thus we obtained many curves of the transient creep on one sample and we could follow the character of the transient creep curves on the applied stress or strain.



a)



b)

Fig.2. The incrementally loading method

The creep measurements have been performed using the equipment designed by the Department of General Physics, Faculty of Science, Masaryk University in Brno. Its working mechanism is schematically illustrated in Fig.3.

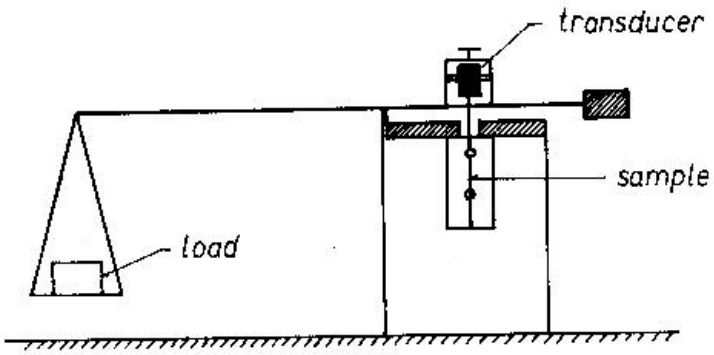


Fig.3. Schematic drawing of the creep apparatus

The loading lever is carried on a hardened steel seat and the leverage ration is 5 : 1. Shots which are added into the vessel attached at the end of the lever are used for loading the lever.

The elongation of the sample is measured by means of a linear variable differential transformer (G.L.Collins Corp., Long Beach, USA). It is situated at the end stretch rod. The range of measurement from the viewpoint of necessary linearity is about ± 0,3 cm. The sensitivity is stated to be about 2V/mm by the DC feeding 6 V from the stabilized power supply Tesla BS 448 E.The linearity of the above mentioned range of measurements is 0,95%. By means of a potentiometer connection a suitable sensitivity for each measurement has been obtained. During the tests the output of the transformer is continuously recorded usány the linear recorder EZ 4 or EZ 11. The whole equipment is situated on a concrete base to ovoid possible vibration.

The temperature was measured by means of a platinum resistor or a Pt-PtRh thermocouple which were placed near the sample or on its surface.

The flow stress of a crystal τ_a can be decomposed into two components τ_i and τ^*. The first one reflects the microstructure (internal long range elastic interactions among obstacles and dislocations). The second component (τ^*) is the stress necessary to push dislocations over local energy barriers (small obstacles, an intrinsic lattice resistance). Then we can write

$$\tau_a = \tau_i + \tau^* \qquad (1)$$

Internal stress τ_i is slowly decreasing with increasing temperature (similarly as elastic constants). Short range interactions of dislocations with energy barriers (described by τ^*) takes place in such a small volume that it is strongly influenced by thermal vibrations. Thermal activation helps dislocation to overcome these barriers thus the flow stress is decreasing with increasing temperature. These short – range thermally activated processes govern almost all the temperature dependent mechanical properties of materials for example dynamic strain aging. The dynamic strain aging occurs usually in the intermediate temperature range (usually 0,3 – 0,4 T_m, where T_m is melting temperature). Among exterior features of dynamic strain aging phenomenon include Portevin – Le Chatelier effect, yield stress plateau and blue brittleness.

2. THERMALLY ACTIVATION THEORY

The concept of thermally activated plastic deformation was introduced as early as 1925 when Becker [1] applied the Boltzmann principle to the nucleation of a slip region. After the introduction of the absolute reaction rate theory of Eyring [2], Kauzmann [3] formulated a general chemical rate theory of plasticity. Similar equations were derived by Seitz and Read [4] based on thermally activated dislocation motion and by Nowick and Machlin [5] based on thermally activated dislocation generation. Later many efforts have been concentrated on definition of activation parameters, its measurement and interpretation [6, 7].

2.1. Activation parameters

The average velocity of a dislocation traveling in an crystal can be considered as a thermally activated process, governed by the Arrhenius type equation

$$v = v_0 \exp\left(-\frac{\Delta F}{kT}\right) \qquad (2)$$

where ΔF is the standard free energy of activation, k is the Boltzmann constant, T is the absolute temperature and v_0 is the velocity when ΔF is zero. The term v_0 may contain the mean distance the dislocation moves per activation event, a fundamental frequency such as kT/h with h being the Planck constant, and a possible geometric factor. On the other hand, v_0 can simply be regarded as the maximum attainable velocity such as shear wave velocity in the crystal.

If a shear stress τ^* is applied in the slip plane so that τ^* does positive work when the dislocation moves forward, then the free energy of activation is decreased for forward motion and increased for backward motion by $\tau^* bA^*$, where b is the Burgers vector of the dislocation and A^* is the area swept by the dislocation during an activation event (activation area). This indicates that external stress may fully activate the dislocation. The stress that can achieve this is τ_c^* which is defined as the friction stress. Let the activation area be A_0^* at $\tau^* = 0$; a consideration of the reversible process shows

$$\Delta F_0 = b\int_0^{A_0^*} \tau^* dA^* \tag{3}$$

Assuming that a relation exists between τ^* and A^* during the activation event. Hence at an applied stress τ^*, the activation free energy for the forward motion is

$$\Delta F_f = \Delta F_0 - b.\tau^* A^* - b\int_{A^*}^{A_0^*} \tau^* dA^* = \Delta F_0 - b\int_0^{\tau^*} A^* d\tau^* \tag{4}$$

Similarly, the activation free energy for backward motion is

$$\Delta F_b = \Delta F_0 + b\int_0^{\tau^*} A^* d\tau^* \tag{5}$$

Equations (2), (4) and (5) give the average velocity of the dislocations [8,9]:

$$v = 2v_c \exp\left(-\frac{\Delta F_0}{k.T}\right)\sinh\frac{b}{k.T}\int_0^{\tau^*} A^* d\tau^* \tag{6}$$

which at small τ^* gives

$$v = 2v_c \frac{A_0^*.\tau^* b}{kT}\exp\left(-\frac{\Delta F_0}{kT}\right) \tag{7}$$

a linear relation between stress and velocity. At large τ^* the velocity becomes

$$v = v_c \exp\left(-\frac{\Delta F_0 - b\int_0^{\tau^*} A^* d\tau^*}{kT} \right) \tag{8}$$

A comparison with (2) shows

$$A^* = -\frac{1}{b}\left(\frac{\partial \Delta F}{\partial \tau^*}\right)_{T,p} = \frac{kT}{b}\left[\frac{\partial \ln\left(\frac{v}{v_c}\right)}{\partial \tau^*}\right]_{T,p} \tag{9}$$

It is to be noted that Eq. (9) is valid only if the hyperbolic sine function in (6) can be approximated by an exponential function.

In the literature the quantity A^*b is sometimes called the "activation volume". To avoid confusion with the activation volume defined as the pressure derivative of the standard free energy of activation, the term "activation area" is defined by Eq. (9).

Similarly other thermodynamic functions can be derived [6,7]:

The „Activation Enthalpy":

$$\Delta H = -b.A^* T\left(\frac{\partial \tau^*}{\partial T}\right)_{p, v/v_c} \tag{10}$$

Activation area of Cd single crystals with various amount of Zn has been measured in wide stress and temperature range (1,5 K – 380 K) [10,11]. The values of A^* and its temperature and stress dependence indicate in the temperature regions ~ 20 K and ~ 200 K changes of mechanisms controlling movement of dislocations.

2.2. The velocity - stress relation.

It was undoubtedly established that activation area decreases with increasing stress [12,13] When the activation area can be approximated by an inverse proportionality to the stress, a velocity – stress relation results:

$$v = B\left(\tau^*\right)^n \tag{11}$$

In this equation *B* and *n* are independent of stress but may be functions of both temperature and pressure. Parameter *n* can be defined by the equation

$$n = \left(\frac{\partial \ln v}{\partial \ln \tau^*} \right)_{T,p} = \frac{\tau^* b A^*}{kT} \tag{12}$$

Equation (12) indicates that temperature dependence *n(T)* can be similar as $A^*(T)$ good qualitative criterion of dislocations mechanisms change (the first one is better because *n* is usually stress independent – Figs.6,7,and 8. [13])..

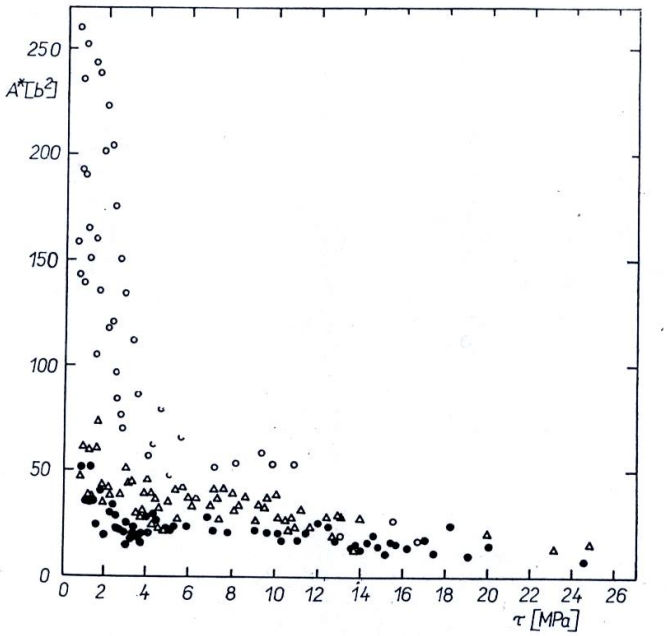


Fig.4. The typical stress dependence of activation area A^* at very low temperatures (Cd+0.0584 at%Zn alloy single crystals deformed at the temperature ◦ - 1.5 K, • - 2,8 K and 4.2 K)

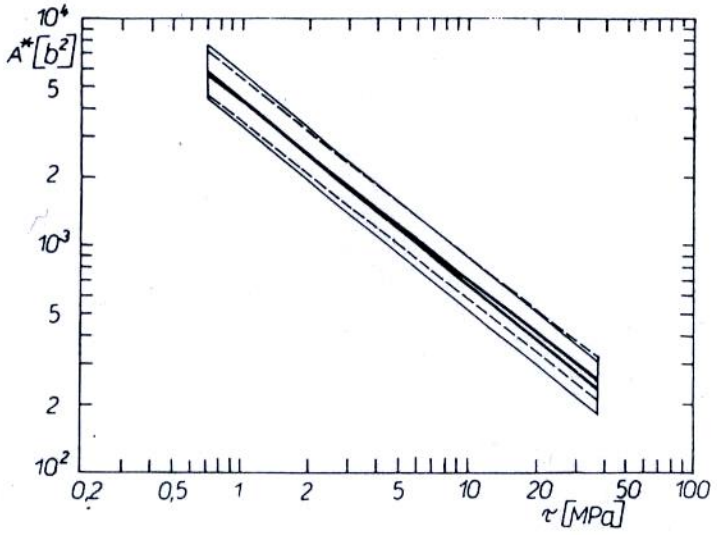


Fig.5 .The stress dependence of the activation area A^*, calculated by means

of the last square method (Cd + 0.0584 at.%Zn, T=77K)
— - The adding of the load increment, ---- the removing of the load increment

The velocity stress exponent *n* for Cd single crystals with various concentration of Zn solute atoms had been measured in the wide temperature interval (1,5 K – 380 K). A repeated creep experiment was used [11,12] and in every creep step one or more values of *n* were measured according to equation (14):

$$n = \left(\frac{\Delta \ln \dot{\varepsilon}_2 / \dot{\varepsilon}_1}{\Delta \ln \tau^*} \right)_{T,p} \qquad (13)$$

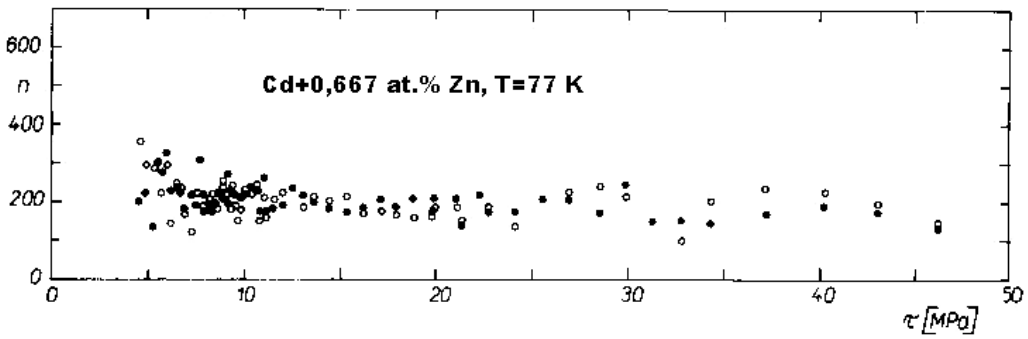


Fig.6. The stress dependence of the stress sensitivity parameter *n*

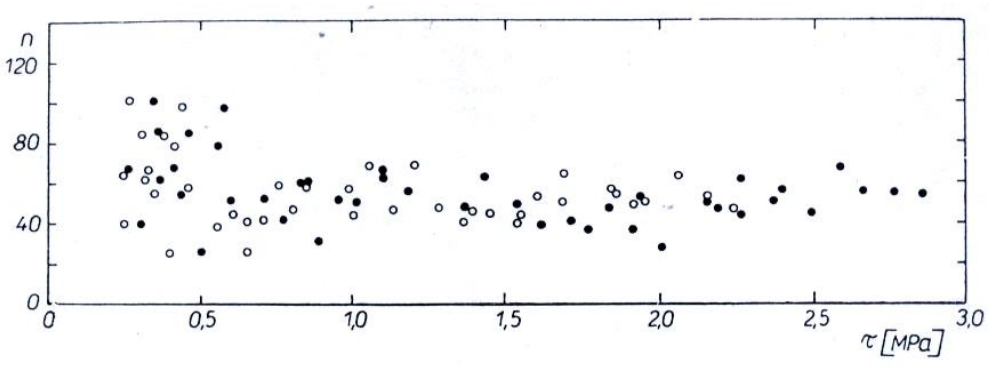


Fig.7.The stress dependence of the stress sensitivity parameter *n* (Cd + 0.0027at.%Zn, T=296K) ◦ - the beginning of the primary creep step, • - the end of the primary creep step.

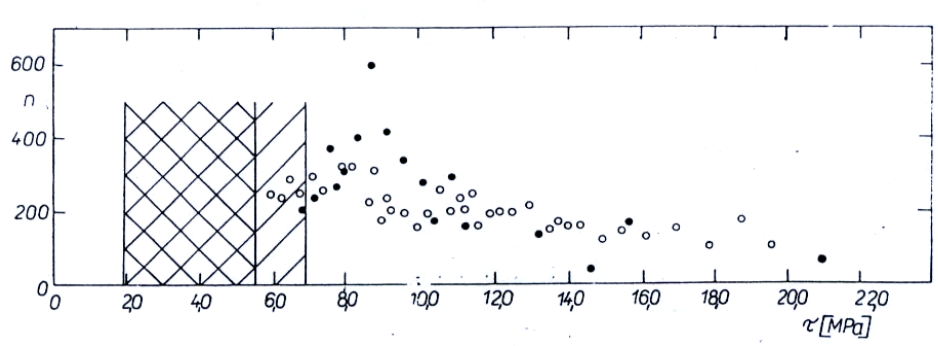


Fig.8.The stress dependence of the stress sensitivity parameter *n* for two samples of the same concentrations of solute atoms, (Cd + 0.0027 at.%Zn, T=202K). ◦ - the sample Cd-Zn 21, • - the sample Cd-Zn 22, ./// the jerky flow (dynamic strain aging)

According to basic equation of plastic deformation (Orowan) we can write

$$\dot{\varepsilon} = b.v.\rho \qquad (14)$$

($\dot{\varepsilon}$ is velocity of deformation and v resp. ρ is velocity resp. density of movable dislocations. We suppose that τ_i does not change in the course of small change of τ_a, i.e $\Delta\,\tau_a \sim \Delta\tau^*$ and $\rho = \text{const}$).

As we can see from the Fig.6,7 and 8., the stress dependence of *n* is approximately constant. The temperature dependence of *n* is for various CdZn alloys shown in the Fig.9.. At that figure we can notice two peaks (at the temperature $T \sim 12$ K and $T \sim 200$ K).

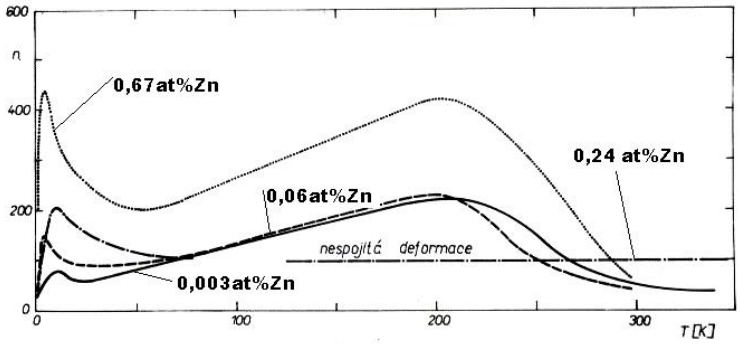


Fig. 9. Temperature dependence of the stress velocity exponent *n.*

3. CONCLUSION

According to our experimental results we can conclude, that

- activation area A^* and the velocity stress exponent *n* can be comparatively easy measured in creep deformation,
- velocity stress exponent *n* is stress independent,
- the temperature dependence $n = n(T)$ indicate changes of mechanisms, controlling velocity of dislocations (at the temperature interval $T \sim 12$ K it is quasidynamical mechanism [14] and at the temperature interval $T \sim 200$ K ($\approx 0{,}3\ T_m$) it is the dynamic strain aging region. [12,13])

REFERENCES

[1] BECKER, R.: *Physikalische Zeitschrift* **26**, 919 (1925)

[2] EYRING, H.J.: *J. Chem. Phys*. **4**, 283, (1936)

[3] KAUZMANN, W.: *Trans. AIME*, **143**, 57, (1941)

[4] SEITZ, F., READ, T.A.: *J. Appl. Phys*. **12**, 100, 170, 470, 538 (1941)

[5] NOVICK, A.S., MACHLUJ, E.S.: *J. Appl. Phys*. 18, 79, (1947)

[6] ROSENFIELD, A.R., HAHN, G.T., BEMENT, A.L., JAFFEE, R.I.: *Dislocation Dynamics*, Mc Graw Hill, 1967.

[7] MARTIN, J.L., CAILLARD, D.: *Thermally Activated Mechanisms in Crystal Plasticity*. Pergamon Press, 2003.

[8] CHRISTIAN, J.W., MASTERS, B.C.,: *Proc. Roy. Soc*. A 281, 240 (1964)

[9] Li, J.C.M.: *Trans. AIME*, **233**, 219 (1965)

[10] HAMERSKÝ, M., NAVRÁTIL, V., LUKÁČ, P., SOLDATOV, V.P. STARTSEV, V.I. *MetallicMaterials*, Bratislava, 3s. 20-26. (1982).

[11] HAMERSKÝ, M., NAVRÁTIL, V., STARTSEV, V.I.: *Czech. J. Phys.*, **B31**, 6 s. (1981)

[12] LUKÁČ, P., STULÍKOVÁ, I., NAVRÁTIL, V.: *Czech. J. Phys.*, **B31**s. pp. 130-134. (1981).

[13] NAVRÁTIL, V., NOVOTNÁ, J.:. Matematika tepelně aktivované plastické deformace kovů. In *Aplimat 6th. Int. Conf.*. 2007. vyd. Bratislava : Faculty of Mechanical Engineering Slovak University of Technology in Bratislava, 2007. od s. 125-130, 6 s.

[14] KAMADA, R., YOSHIZAWA,I.: *J.Phys Soc.Japan **31***, 1056-1068, (1971).

Weakly Delayed Systems of Linear Discrete Equations in $\mathbb{R}^3$

Jan Šafařík

Faculty of Civil Engineering,
Faculty of Electrical Engineering and Communication,
Brno University of Technology, Brno, Czech Republic.
safarik.j@fce.vutbr.cz

Josef Diblík

Faculty of Civil Engineering,
Faculty of Electrical Engineering and Communication,
Brno University of Technology, Brno, Czech Republic.
diblik.j@fce.vutbr.cz

Hana Halfarová

Faculty of Civil Engineering, Brno University of Technology,
Brno, Czech Republic.
halfarova.h@fce.vutbr.cz

Abstract: *The purpose of this paper is to provide criteria for a linear discrete system in $\mathbb{R}^3$ with delay to be weakly delayed. Explicit necessary and sufficient conditions are derived.*

Keywords: Discrete linear system, weakly delayed system, delay.

1 Weakly Delayed Systems

We use the following notation throughout this paper: For integers s, q, $s \leq q$, we define a set $\mathbb{Z}_s^q := \{s, s+1, \dots, q-1, q\}$. Similarly, we define a set $\mathbb{Z}_s^\infty := \{s, s+1, \dots\}$. In this paper, we deal with the discrete systems

$$x(k+1) = Ax(k) + Bx(k-m) \tag{1}$$

where $m \geq 0$ are fixed integers, $k \in \mathbb{Z}_0^\infty$, $A = (a_{ij})$ and $B = (b_{ij})$, are constant $l \times l$ matrices, and $x\colon \mathbb{Z}_{-m}^\infty \to \mathbb{R}^l$, $l \geq 2$.

In [2], linear weakly delayed systems were defined for planar systems. This definition can be applied to l-dimensional systems as follows.

Definition 1 *System* (1) *is called weakly delayed if the characteristic equations for* (1) *and for the system without delay*

$$x(k+1) = Ax(k)$$

have identical roots, that is, if, for every $\lambda \in \mathbb{C} \setminus \{0\}$,

$$\det\left(A + \lambda^{-m}B - \lambda I\right) = \det\left(A - \lambda I\right).$$

2 Criteria of Weakly Delayed Systems

In [2], the authors derive necessary and sufficient conditions for (1) with $l = 2$ to be a weakly delayed system:

Theorem 1 *System (1) is a system with weak delay if and only if the following three conditions hold simultaneously:*

$$b_{11} + b_{22} = 0,$$

$$\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = 0,$$

$$\begin{vmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} \\ a_{21} & a_{22} \end{vmatrix} = 0.$$

Moreover, in [4] Theorem 1 is extended to the case $l = 3$.

Theorem 2 ([4]) *Let $l = 3$ in (1). Then, (1) is a weakly delayed system if and only if conditions (2)–(7) below hold:*

$$b_{11} + b_{22} + b_{33} = 0, \tag{2}$$

$$\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} = 0, \tag{3}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0, \tag{4}$$

$$\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} = 0, \tag{5}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0, \tag{6}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0. \tag{7}$$

In the following part of the paper, considering equation (1) with $l = 3$, we will simplify conditions (2)–(7) for every possible Jordan canonical form of the matrix A.

3 Jordan Canonical Forms of A and Criteria for Weakly Delayed Systems

It is known that, for every matrix A, there exists a nonsingular matrix S transforming it to the corresponding Jordan matrix form A^*. This means that

$$A^* = S^{-1}AS$$

where A^* has the following seven possible forms (denoted below as $A_1, \ldots, A_7$), depending on the roots of the characteristic equation

$$\det(A - \lambda I) = 0. \tag{8}$$

Throughout the remaining part of the paper we assume that $l = 3$ in (1).
If (8) has three real distinct roots λ_1, λ_2, λ_3, then

$$A_1 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \tag{9}$$

if (8) has one double real root λ_1, $\lambda_2 = \lambda_3$, then

$$A_2 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} \tag{10}$$

or

$$A_3 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix}, \tag{11}$$

in the case of one triple real root $\lambda = \lambda_{1,2,3}$, the following forms are possible

$$A_4 = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \tag{12}$$

$$A_5 = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \tag{13}$$

$$A_6 = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \tag{14}$$

and, finally, if one root is real and two roots are complex conjugate, i.e. $\lambda_{2,3} = p \pm iq$, with $q \neq 0$, then

$$A_7 = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & p & q \\ 0 & -q & p \end{pmatrix}. \tag{15}$$

In this part, we will simplify the general conditions (2)–(7) for each of the Jordan forms (9)–(15).

3.1 Criterion for Weakly Delayed Systems in the Case (9)

Consider system (1) with the matrix $A = A_1$, i.e.,

$$x(k+1) = A_1 x(k) + Bx(k-m). \tag{16}$$

In [3] the following result is formulated.

Theorem 3 *System (16) is a weakly delayed system if and only if*

$$b_{11} = b_{22} = b_{33} = 0, \tag{17}$$
$$b_{12}b_{23}b_{31} + b_{13}b_{21}b_{32} = 0, \tag{18}$$
$$b_{12}b_{21} + b_{13}b_{31} + b_{23}b_{32} = 0, \tag{19}$$
$$\lambda_3 b_{12}b_{21} + \lambda_2 b_{13}b_{31} + \lambda_1 b_{23}b_{32} = 0. \tag{20}$$

We will show the proof of Theorem 3 as it is not given in [3].

Proof. It is possible to simplify conditions (4), (6) and (7). From (4) we get

$$\lambda_1(b_{22}b_{33} - b_{23}b_{32}) + \lambda_2(b_{11}b_{33} - b_{13}b_{31}) + \lambda_3(b_{11}b_{22} - b_{12}b_{21}) = 0$$

because

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$
$$= \begin{vmatrix} \lambda_1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda_2 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda_3 \end{vmatrix} =$$
$$= \lambda_1(b_{22}b_{33} - b_{23}b_{32}) + \lambda_2(b_{11}b_{33} - b_{13}b_{31}) + \lambda_3(b_{11}b_{22} - b_{12}b_{21}) =$$
$$= 0.$$

From (6) we get

$$\lambda_2\lambda_3 b_{11} + \lambda_1\lambda_3 b_{22} + \lambda_1\lambda_2 b_{33} = 0 \tag{21}$$

since

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$
$$= \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} \lambda_1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda_3 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{vmatrix} =$$
$$= \lambda_2\lambda_3 b_{11} + \lambda_1\lambda_3 b_{22} + \lambda_1\lambda_2 b_{33} = 0.$$

From (7) we get

$$(\lambda_2 + \lambda_3)b_{11} + (\lambda_1 + \lambda_3)b_{22} + (\lambda_1 + \lambda_2)b_{33} = 0 \tag{22}$$

since

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$
$$+ \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$
$$= \begin{vmatrix} \lambda_1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$
$$+ \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ 0 & 0 & \lambda_3 \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda_3 \end{vmatrix} =$$
$$= (\lambda_2 + \lambda_3)b_{11} + (\lambda_1 + \lambda_3)b_{22} + (\lambda_1 + \lambda_2)b_{33} = 0.$$

From (2), (21) and (22) we deduce

$$\begin{array}{rcrcrcl} b_{11} & + & b_{22} & + & b_{33} & = & 0, \\ \lambda_2\lambda_3 b_{11} & + & \lambda_1\lambda_3 b_{22} & + & \lambda_1\lambda_2 b_{33} & = & 0, \\ (\lambda_2 + \lambda_3)b_{11} & + & (\lambda_1 + \lambda_3)b_{22} & + & (\lambda_1 + \lambda_2)b_{33} & = & 0. \end{array} \tag{23}$$

The determinant of the system (23) is different from zero since

$$\begin{vmatrix} 1 & 1 & 1 \\ \lambda_2\lambda_3 & \lambda_1\lambda_3 & \lambda_1\lambda_2 \\ (\lambda_2 + \lambda_3) & (\lambda_1 + \lambda_3) & (\lambda_1 + \lambda_2) \end{vmatrix} = \lambda_1\lambda_3(\lambda_1 + \lambda_2) - \lambda_1\lambda_2(\lambda_1 + \lambda_3) -$$
$$-\lambda_2\lambda_3(\lambda_1 + \lambda_2) + \lambda_1\lambda_2(\lambda_2 + \lambda_3) + \lambda_2\lambda_2(\lambda_1 + \lambda_3) - \lambda_1\lambda_3(\lambda_2 + \lambda_3) =$$
$$= -(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3) \neq 0.$$

Consequently, (23) has only the trivial solution

$$b_{11} = b_{22} = b_{33} = 0. \tag{24}$$

Therefore,

$$B = \begin{pmatrix} 0 & b_{12} & b_{13} \\ b_{21} & 0 & b_{23} \\ b_{31} & b_{32} & 0 \end{pmatrix}.$$

Applying (24) to (3)–(7), after simplification, we get (18)–(20).

Example 1 *Assume that* $A_1 = \operatorname{diag}(0, 1, 2)$,

$$B = \begin{pmatrix} 0 & -1 & 2 \\ -2 & 0 & 2 \\ -2 & 1 & 0 \end{pmatrix}.$$

It is easy to verify that conditions (17)–(20) *are valid and system* (16) *is weakly delayed.*

3.2 Criterion for Weakly Delayed Systems in the Case (10)

Consider system (1) with the matrix $A = A_2$, i.e.,

$$x(k+1) = A_2 x(k) + Bx(k-m). \tag{25}$$

Theorem 4 *System (25) is a weakly delayed system if and only if*

$$b_{11} = 0, \tag{26}$$
$$b_{22} + b_{33} = 0, \tag{27}$$
$$b_{12}b_{21} + b_{13}b_{31} = 0, \tag{28}$$
$$b_{22}b_{33} + b_{23}b_{32} = 0, \tag{29}$$
$$b_{12}b_{23}b_{31} + b_{13}b_{21}b_{32} - b_{13}b_{22}b_{31} - b_{12}b_{21}b_{33} = 0. \tag{30}$$

Proof. It is possible to simplify conditions (4), (6) and (7). From (4) we get

$$\lambda_1(b_{22}b_{33} - b_{23}b_{32}) + \lambda_2(b_{11}b_{22} + b_{11}b_{33} - b_{12}b_{21} - b_{13}b_{31}) = 0 \tag{31}$$

because

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$
$$= \begin{vmatrix} \lambda_1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda_2 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda_2 \end{vmatrix} =$$
$$= \lambda_1(b_{22}b_{33} - b_{23}b_{32}) + \lambda_2(b_{11}b_{33} - b_{13}b_{31}) + \lambda_2(b_{11}b_{22} - b_{12}b_{21}) =$$
$$= \lambda_1(b_{22}b_{33} - b_{23}b_{32}) + \lambda_2(b_{11}b_{22} + b_{11}b_{33} - b_{12}b_{21} - b_{13}b_{31}) = 0.$$

From (6) we get

$$\lambda_2^2 b_{11} + \lambda_1\lambda_2(b_{22} + b_{33}) = 0 \tag{32}$$

since

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$
$$= \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} \lambda_1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda_2 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_2 \end{vmatrix} =$$
$$= \lambda_1\lambda_2 b_{33} + \lambda_1\lambda_2 b_{22} + \lambda_2^2 b_{11} = \lambda_2^2 b_{11} + \lambda_1\lambda_2(b_{22} + b_{33}) = 0.$$

From (7) we get

$$\lambda_1(b_{22} + b_{33}) + \lambda_2(2b_{11} + b_{22} + b_{33}) = 0 \tag{33}$$

since

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$
$$+ \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$
$$= \begin{vmatrix} \lambda_1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$
$$+ \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ 0 & 0 & \lambda_2 \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda_2 \end{vmatrix} =$$
$$= \lambda_1 b_{22} + \lambda_1 b_{33} + \lambda_2 b_{33} + \lambda_2 b_{11} + \lambda_2 b_{11} + \lambda_2 b_{22} =$$
$$= \lambda_1 (b_{22} + b_{33}) + \lambda_2 (2b_{11} + b_{22} + b_{33}) = 0.$$

From (2) we get $b_{22} + b_{33} = -b_{11}$ and (32) yields

$$\begin{aligned} \lambda_2^2 b_{11} + \lambda_1 \lambda_2 (b_{22} + b_{33}) &= 0, \\ \lambda_2^2 b_{11} + \lambda_1 \lambda_2 (-b_{11}) &= 0, \\ b_{11} \lambda_2 (\lambda_2 - \lambda_1) &= 0. \end{aligned}$$

Since $\lambda_2 - \lambda_1 \neq 0$, we get

$$\lambda_2 b_{11} = 0$$

Assume $\lambda_2 = 0$. Then, from (33) we have $\lambda_1 (b_{22} + b_{33}) = 0 \Rightarrow \lambda_1 b_{11} = 0$. Because $\lambda_1 \neq 0$, we get

$$b_{11} = 0. \tag{34}$$

From (2), using (34), we have

$$b_{22} + b_{33} = 0.$$

From (31) we get

$$b_{22} b_{33} - b_{23} b_{23} = 0. \tag{35}$$

Substitute (34) and (35) into (5). We get

$$b_{12} b_{21} + b_{13} b_{31} = 0.$$

If $\lambda_2 \neq 0$, we get the same conditions (26)–(29). Condition (3) can be simplified to (30).

3.3 Criterion for Weakly Delayed Systems in the Case (11)

Consider system (1) with the matrix $A = A_3$, i.e.,

$$x(k+1) = A_3 x(k) + B x(k-m). \tag{36}$$

Theorem 5 *System* (36) *is a weakly delayed system if and only if*

$$\begin{aligned} b_{11} &= 0, \\ b_{22} + b_{33} &= 0, \\ b_{32} &= 0, \\ (\lambda_1 - \lambda_2) b_{22} b_{33} + b_{12} b_{31} &= 0, \\ b_{12} b_{23} b_{31} - b_{13} b_{22} b_{31} - b_{12} b_{21} b_{33} &= 0. \end{aligned} \tag{37}$$

Proof. It is possible to simplify conditions (4), (6) and (7). From (4) we get

$$\begin{aligned} \lambda_1 (b_{22} b_{33} - b_{23} b_{32}) + \lambda_2 (b_{11} b_{22} + b_{11} b_{33} - b_{12} b_{21} - b_{13} b_{31}) \\ - (b_{11} b_{32} - b_{12} b_{31}) = 0 \end{aligned} \tag{38}$$

because

$$\begin{aligned} &\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \\ &= \begin{vmatrix} \lambda_1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda_2 & 1 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda_2 \end{vmatrix} = \\ &= \lambda_1 (b_{22} b_{33} - b_{23} b_{32}) + \lambda_2 (b_{11} b_{33} - b_{13} b_{31}) - (b_{11} b_{32} - b_{12} b_{31}) \\ &\quad + \lambda_2 (b_{11} b_{22} - b_{12} b_{21}) = \\ &= \lambda_1 (b_{22} b_{33} - b_{23} b_{32}) + \lambda_2 (b_{11} b_{22} + b_{11} b_{33} - b_{12} b_{21} - b_{13} b_{31}) \\ &\quad - (b_{11} b_{32} - b_{12} b_{31}) = 0. \end{aligned}$$

From (6) we get

$$\lambda_2^2 b_{11} + \lambda_1 \lambda_2 (b_{22} + b_{33}) - \lambda_1 b_{32} = 0 \tag{39}$$

since

$$\begin{aligned} &\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \\ &= \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} \lambda_1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda_2 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{vmatrix} = \\ &= \lambda_1 \lambda_2 b_{33} - \lambda_1 b_{32} + \lambda_1 \lambda_2 b_{22} + \lambda_2^2 b_{11} = \\ &= \lambda_2^2 b_{11} + \lambda_1 \lambda_2 (b_{22} + b_{33}) - \lambda_1 b_{32} = 0. \end{aligned}$$

From (7) we get

$$\lambda_1 (b_{22} + b_{33}) + \lambda_2 (2 b_{11} + b_{22} + b_{33}) - b_{32} = 0 \tag{40}$$

since

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$

$$
\begin{aligned}
&+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \\
&= \begin{vmatrix} \lambda_1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} \\
&+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda_2 & 1 \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ 0 & 0 & \lambda_2 \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda_2 \end{vmatrix} = \\
&=\lambda_1 b_{22} + \lambda_1 b_{33} + \lambda_2 b_{33} - b_{32} + \lambda_2 b_{11} + \lambda_2 b_{11} + \lambda_2 b_{22} = \\
&=\lambda_1(b_{22} + b_{33}) + \lambda_2(2b_{11} + b_{22} + b_{33}) - b_{32} = 0.
\end{aligned}
$$

From (2) we have $b_{22} + b_{33} = -b_{11}$. Moreover, from (40) we get

$$
\begin{aligned}
\lambda_1(b_{22} + b_{33}) + \lambda_2(2b_{11} + b_{22} + b_{33}) - b_{32} &= 0, \\
\lambda_1(-b_{11}) + \lambda_2 b_{11} - b_{32} &= 0, \\
b_{11}(\lambda_2 - \lambda_1) - b_{32} &= 0.
\end{aligned} \tag{41}
$$

From the last expression, we obtain $b_{32} = b_{11}(\lambda_2 - \lambda_1)$. Substituting it into (39), we have

$$
\begin{aligned}
\lambda_2^2 b_{11} + \lambda_1\lambda_2(-b_{11}) - \lambda_1 b_{11}(\lambda_2 - \lambda_1) &= 0, \\
b_{11}(\lambda_1^2 - 2\lambda_1\lambda_2 + \lambda_2^2) &= 0, \\
b_{11}(\lambda_1 - \lambda_2)^2 &= 0.
\end{aligned}
$$

Because $(\lambda_1 - \lambda_2)^2 \neq 0$, we derive

$$
b_{11} = 0. \tag{42}
$$

From (2), utilizing (42), we have

$$
b_{22} + b_{33} = 0.
$$

Similarly, from (41), using (42), we obtain

$$
b_{32} = 0. \tag{43}
$$

Substitute (42) and (43) into (5). We get

$$
-b_{12}b_{21} - b_{13}b_{31} + b_{22}b_{33} = 0 \Rightarrow b_{22}b_{33} = b_{12}b_{21} + b_{13}b_{31}.
$$

Substituting the last expression into (38) we get

$$
\begin{aligned}
\lambda_1(b_{22}b_{33}) + \lambda_2(-b_{12}b_{21} - b_{13}b_{31}) + b_{12}b_{31} &= 0, \\
\lambda_1(b_{22}b_{33}) + \lambda_2(-b_{22}b_{33}) + b_{12}b_{31} &= 0, \\
(\lambda_1 - \lambda_2)b_{22}b_{33} + b_{12}b_{31} &= 0.
\end{aligned}
$$

Condition (3) can easily be simplified to (37).

3.4 Criterion for Weakly Delayed Systems in the Case (12)

Consider system (1) with the matrix $A = A_4$, i.e.,

$$x(k+1) = A_4 x(k) + Bx(k-m). \tag{44}$$

Theorem 6 *System (44) is a weakly delayed system if and only if*

$$\begin{aligned} b_{11} + b_{22} + b_{33} &= 0, \\ b_{11}b_{22} + b_{11}b_{33} + b_{22}b_{33} - b_{12}b_{21} - b_{13}b_{31} - b_{23}b_{32} &= 0, \\ b_{11}b_{22}b_{33} + b_{12}b_{23}b_{31} + b_{13}b_{21}b_{32} - b_{13}b_{22}b_{31} - b_{12}b_{21}b_{33} - b_{11}b_{23}b_{32} &= 0. \end{aligned} \tag{45}$$

Proof. It is possible to simplify conditions (4), (6) and (7). From (4) we get

$$\lambda(b_{11}b_{22} + b_{11}b_{33} + b_{22}b_{33} - b_{12}b_{21} - b_{23}b_{32} - b_{13}b_{31}) = 0 \tag{46}$$

because

$$\begin{aligned} &\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \\ &= \begin{vmatrix} \lambda & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda \end{vmatrix} = \\ &= \lambda(b_{22}b_{33} - b_{23}b_{32}) + \lambda(b_{11}b_{33} - b_{13}b_{31}) + \lambda(b_{11}b_{22} - b_{12}b_{21}) = \\ &= \lambda(b_{11}b_{22} + b_{11}b_{33} + b_{22}b_{33} - b_{12}b_{21} - b_{23}b_{32} - b_{13}b_{31}) = 0. \end{aligned}$$

From (6) we get

$$\lambda^2(b_{11} + b_{22} + b_{33}) = 0 \tag{47}$$

since

$$\begin{aligned} &\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \\ &= \begin{vmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} \lambda & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{vmatrix} = \\ &= \lambda^2 b_{33} + \lambda^2 b_{22} + \lambda^2 b_{11} = \lambda^2(b_{11} + b_{22} + b_{33}) = 0. \end{aligned}$$

From (7) we get

$$2\lambda(b_{11} + b_{22} + b_{33}) = 0 \tag{48}$$

since

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$

$$+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= \begin{vmatrix} \lambda & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$

$$+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda \end{vmatrix} =$$

$$= \lambda b_{22} + \lambda b_{33} + \lambda b_{33} + \lambda b_{11} + \lambda b_{11} + \lambda b_{22} =$$

$$= 2\lambda(b_{11} + b_{22} + b_{33}) = 0.$$

It is easy to see that (47) as well as (48) are valid because (2) holds, (46) is valid because (5) holds. Condition (45) is equivalent to (3).

3.5 Criterion for Weakly Delayed Systems in the Case (13)

Consider system (1) with the matrix $A = A_5$, i.e.,

$$x(k+1) = A_5 x(k) + Bx(k-m). \tag{49}$$

Theorem 7 *System* (49) *is a weakly delayed system if and only if*

$$\begin{aligned} b_{11} + b_{22} + b_{33} &= 0, \\ b_{21} &= 0, \\ b_{23} b_{31} &= 0, \\ b_{11} b_{22} + b_{11} b_{33} + b_{22} b_{33} - b_{13} b_{31} - b_{23} b_{32} &= 0, \qquad (50) \\ b_{11} b_{22} b_{33} - b_{13} b_{22} b_{31} - b_{11} b_{23} b_{32} &= 0. \qquad (51) \end{aligned}$$

Proof. It is possible to simplify conditions (4), (6) and (7). From (4) we get

$$\lambda(b_{11}b_{22} + b_{11}b_{33} + b_{22}b_{33} - b_{12}b_{21} - b_{13}b_{31} - b_{23}b_{32}) - (b_{21}b_{33} - b_{23}b_{31}) = 0 \tag{52}$$

because

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= \begin{vmatrix} \lambda & 1 & 0 \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda \end{vmatrix} =$$

$$= \lambda(b_{22}b_{33} - b_{23}b_{32}) - (b_{21}b_{33} - b_{23}b_{31}) + \lambda(b_{11}b_{33} - b_{13}b_{31})$$

$$
\begin{aligned}
&+\lambda(b_{11}b_{22}-b_{12}b_{21})= \\
=&\lambda(b_{11}b_{22}+b_{11}b_{33}+b_{22}b_{33}-b_{12}b_{21}-b_{13}b_{31}-b_{23}b_{32}) \\
&-(b_{21}b_{33}-b_{23}b_{31})=0.
\end{aligned}
$$

From (6) we get

$$\lambda^2(b_{11}+b_{22}+b_{33})-\lambda b_{21}=0 \tag{53}$$

since

$$
\begin{aligned}
&\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}+\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}= \\
=&\begin{vmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix}+\begin{vmatrix} \lambda & 1 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda \end{vmatrix}+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{vmatrix}= \\
=&\lambda^2 b_{33}+\lambda^2 b_{22}-\lambda b_{21}+\lambda^2 b_{11}=\lambda^2(b_{11}+b_{22}+b_{33})-\lambda b_{21}=0.
\end{aligned}
$$

From (7) we get

$$2\lambda(b_{11}+b_{22}+b_{33})-b_{21}=0 \tag{54}$$

since

$$
\begin{aligned}
&\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix}+\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix}+\begin{vmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} \\
&+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{vmatrix}+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix}+\begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}= \\
=&\begin{vmatrix} \lambda & 1 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix}+\begin{vmatrix} \lambda & 1 & 0 \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix}+\begin{vmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} \\
&+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{vmatrix}+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{vmatrix}+\begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda \end{vmatrix}= \\
=&\lambda b_{22}-b_{21}+\lambda b_{33}+\lambda b_{33}+\lambda b_{11}+\lambda b_{11}+\lambda b_{22}= \\
=&2\lambda(b_{11}+b_{22}+b_{33})-b_{21}=0.
\end{aligned}
$$

If (2) holds, we get

$$b_{21}=0. \tag{55}$$

Moreover, (53) holds because of (54). From (5) and (55), then, (52) yields

$$b_{23}b_{31}=0.$$

Equation (5) can be simplified to (50). Condition (51) can be obtained from (3) using the equation derived above.

3.6 Criterion for Weakly Delayed Systems in the Case (14)

Consider system (1) with the matrix $A = A_6$, i.e.,

$$x(k+1) = A_6 x(k) + Bx(k-m). \tag{56}$$

Theorem 8 *System (56) is a weakly delayed system if and only if*

$$\begin{aligned}
b_{11} + b_{22} + b_{33} &= 0, \\
b_{21} + b_{32} &= 0, \\
b_{31} &= 0, \\
b_{21}b_{33} + b_{11}b_{32} &= 0, && (57)\\
b_{11}b_{22} + b_{11}b_{33} + b_{22}b_{33} - b_{12}b_{21} - b_{23}b_{32} &= 0, && (58)\\
b_{11}b_{22}b_{33} + b_{13}b_{21}b_{32} - b_{12}b_{21}b_{33} - b_{11}b_{23}b_{32} &= 0. && (59)
\end{aligned}$$

Proof. It is possible to simplify conditions (4), (6) and (7). From (4) we get

$$\begin{aligned}
&\lambda(b_{11}b_{22} + b_{11}b_{33} + b_{22}b_{33} - b_{12}b_{21} - b_{13}b_{31} - b_{23}b_{32}) \\
&\qquad - (b_{11}b_{32} + b_{21}b_{33} - b_{12}b_{31} - b_{23}b_{31}) = 0
\end{aligned} \tag{60}$$

because

$$\begin{aligned}
&\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \\
&= \begin{vmatrix} \lambda & 1 & 0 \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda & 1 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda \end{vmatrix} = \\
&= \lambda(b_{22}b_{33} - b_{23}b_{32}) - (b_{21}b_{33} - b_{23}b_{31}) + \lambda(b_{11}b_{33} - b_{13}b_{31}) \\
&\quad - (b_{11}b_{32} - b_{12}b_{31}) + \lambda(b_{11}b_{22} - b_{12}b_{21}) = \\
&= \lambda(b_{11}b_{22} + b_{11}b_{33} + b_{22}b_{33} - b_{12}b_{21} - b_{13}b_{31} - b_{23}b_{32}) \\
&\quad - (b_{11}b_{32} + b_{21}b_{33} - b_{12}b_{31} - b_{23}b_{31}) = 0.
\end{aligned}$$

From (6) we get

$$\lambda^2(b_{11} + b_{22} + b_{33}) - \lambda(b_{21} + b_{32}) + b_{31} = 0 \tag{61}$$

since

$$\begin{aligned}
&\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \\
&= \begin{vmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} \lambda & 1 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{vmatrix} = \\
&= \lambda^2 b_{33} - \lambda b_{32} + b_{31} + \lambda^2 b_{22} - \lambda b_{21} + \lambda^2 b_{11}
\end{aligned}$$

$$=\lambda^2(b_{11}+b_{22}+b_{33})-\lambda(b_{21}+b_{21})+b_{31}=0.$$

From (7) we get

$$2\lambda(b_{11}+b_{22}+b_{33})-(b_{21}+b_{32})=0 \tag{62}$$

since

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$
$$+ \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$
$$= \begin{vmatrix} \lambda & 1 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} \lambda & 1 & 0 \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ 0 & \lambda & 1 \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$
$$+ \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & \lambda & 1 \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & \lambda \end{vmatrix} =$$
$$=\lambda b_{22}-b_{21}+\lambda b_{33}+\lambda b_{33}-b_{32}+\lambda b_{11}+\lambda b_{11}+\lambda b_{22}=$$
$$=2\lambda(b_{11}+b_{22}+b_{33})-(b_{21}+b_{32})=0.$$

If (2) holds, we get

$$b_{21}+b_{32}=0 \tag{63}$$

from (62). Moreover, from (61) we get

$$b_{31}=0 \tag{64}$$

using (2) and (63). From (5) and (64), then, (60) yields (57). Equation (5) can be simplified to (58), equation (3) can be simplified to (59).

3.7 Criterion for Weakly Delayed Systems in the Case (15)

Consider system (1) with the matrix $A=A_6$, i.e.,

$$x(k+1)=A_7x(k)+Bx(k-m). \tag{65}$$

Theorem 9 *System (65) is a weakly delayed system if and only if*

$$\begin{aligned} b_{11}&=0,\\ b_{22}+b_{33}&=0,\\ b_{23}-b_{32}&=0,\\ b_{22}b_{33}-b_{12}b_{21}-b_{13}b_{31}-b_{23}b_{32}&=0,\\ (\lambda-p)(b_{12}b_{21}+b_{13}b_{31})+q(b_{12}b_{31}-b_{13}b_{21})&=0,\\ b_{12}b_{23}b_{31}+b_{13}b_{21}b_{32}-b_{13}b_{22}b_{31}-b_{12}b_{21}b_{33}&=0. \end{aligned} \tag{66}$$

Proof. It is possible to simplify conditions (4), (6) and (7). From (4) we get

$$\lambda(b_{22}b_{33} - b_{23}b_{32}) + p(b_{11}b_{22} + b_{11}b_{33} - b_{12}b_{21} - b_{13}b_{31}) + q(b_{11}b_{23} + b_{12}b_{31} - b_{11}b_{32} - b_{13}b_{21}) = 0 \quad (67)$$

because

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$
$$= \begin{vmatrix} \lambda & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & p & q \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & -q & p \end{vmatrix} =$$
$$= \lambda(b_{22}b_{33} - b_{23}b_{32}) + p(b_{11}b_{33} - b_{13}b_{31}) - q(b_{11}b_{32} - b_{12}b_{31}) + q(b_{11}b_{23} - b_{13}b_{21}) + p(b_{11}b_{22} - b_{12}b_{21}) =$$
$$= \lambda(b_{22}b_{33} - b_{23}b_{32}) + p(b_{11}b_{22} + b_{11}b_{33} - b_{12}b_{21} - b_{13}b_{31}) + q(b_{11}b_{23} + b_{12}b_{31} - b_{11}b_{32} - b_{13}b_{21}) = 0.$$

From (6) we get

$$\lambda(p(b_{22} + b_{33}) + q(b_{23} - b_{32})) + b_{11}(p^2 + q^2) = 0 \quad (68)$$

since

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$
$$= \begin{vmatrix} \lambda & 0 & 0 \\ 0 & p & q \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} \lambda & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & -q & p \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & p & q \\ 0 & -q & p \end{vmatrix} =$$
$$= \lambda(pb_{33} - qb_{32}) + \lambda(pb_{22} + qb_{23}) + b_{11}(p^2 + q^2) =$$
$$= \lambda(p(b_{22} + b_{33}) + q(b_{23} - b_{32})) + b_{11}(p^2 + q^2) = 0$$

From (7) we get

$$\lambda(b_{22} + b_{33}) + p(2b_{11} + b_{22} + b_{33}) + q(b_{23} - b_{32}) = 0 \quad (69)$$

since

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$
$$+ \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$
$$= \begin{vmatrix} \lambda & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ 0 & p & q \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$

$$+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & p & q \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 \\ 0 & -q & p \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ 0 & -q & p \end{vmatrix} =$$
$$=\lambda b_{22} + \lambda b_{33} + pb_{33} - qb_{32} + pb_{11} + pb_{11} + pb_{22} + qb_{23} =$$
$$=\lambda(b_{22} + b_{33}) + p(2b_{11} + b_{22} + b_{33}) + q(b_{23} - b_{32}) = 0.$$

From (2), we have $b_{22} + b_{33} = -b_{11}$. Expression (69) yields

$$\begin{aligned} \lambda(b_{22} + b_{33}) + p(2b_{11} + b_{22} + b_{33}) + q(b_{23} - b_{32}) &= 0, \\ \lambda(-b_{11}) + pb_{11} - q(b_{23} - b_{32}) &= 0, \\ -b_{11}(\lambda - p) - q(b_{23} - b_{32}) &= 0. \end{aligned}$$

From the last expression we have $q(b_{23} - b_{32}) = (\lambda - p)b_{11}$. A substitution into (68) yields

$$\begin{aligned} \lambda(p(b_{22} + b_{33}) + q(b_{23} - b_{32})) + b_{11}(p^2 + q^2) &= 0, \\ \lambda p(-b_{11}) + \lambda b_{11}(\lambda - p) + b_{11}(p^2 + q^2) &= 0, \\ b_{11}(\lambda^2 - 2\lambda p + p^2 + q^2) &= 0, \\ b_{11}((\lambda - p)^2 + q^2)^2 &= 0. \end{aligned}$$

Since $((\lambda - p)^2 + q^2)^2 \neq 0$, we get

$$b_{11} = 0. \tag{70}$$

From (2), utilizing (70), we derive

$$b_{22} + b_{33} = 0. \tag{71}$$

Substituting (70) and (71) into (68), we have

$$b_{23} - b_{32} = 0.$$

Simplifying (5) leads to

$$\begin{aligned} (b_{11}b_{22} - b_{12}b_{21}) + (b_{11}b_{33} - b_{13}b_{31}) + (b_{22}b_{33} - b_{23}b_{32}) &= 0, \\ -b_{12}b_{21} - b_{13}b_{31} + b_{22}b_{33} - b_{23}b_{32} &= 0. \end{aligned}$$

Then, from the last expression, we get

$$b_{22}b_{33} - b_{23}b_{32} = b_{12}b_{21} + b_{13}b_{31}.$$

Substituting it together with (70) into (67), we obtain

$$\begin{aligned} \lambda(b_{12}b_{21} + b_{13}b_{31}) + p(-b_{12}b_{21} - b_{13}b_{31}) + q(b_{12}b_{31} - b_{13}b_{21}) &= 0, \\ (b_{12}b_{21} + b_{13}b_{31})(\lambda - p) + q(b_{12}b_{31} - b_{13}b_{21}) &= 0. \end{aligned}$$

Condition (3) can be simplified to (66).

4 Conclusion

Weakly delayed three-dimensional systems of linear discrete equations with constant coefficients and constant delays were considered and criteria for systems (1), with $l = 3$, to be weakly delayed were derived. It is an open question how to derive criteria for systems with several delays. For further results related to weakly delayed systems, we refer, e.g., to [1, 5, 6]

Acknowledgements

The first two authors were supported by the Grant FEKT-S-14-2200 of Faculty of Electrical Engineering and Communication, BUT.

Reference

[1] KHUSAINOV D. Ya., BENDITKIS D. B., Diblík J.: *Weak delay in systems with an aftereffect.* Functional Differential Equations, 9 (2002), pp. 385-404.

[2] Diblík J., Khusainov D. Ya., Šmarda Z.: *Construction of the general solution of planar linear discrete systems with constant coefficients and weak delay*. Adv. Difference Equ. 2009, Art. ID 784935, 18 pp.

[3] DIBLÍK J., HALFAROVÁ H.: *Linear system of discrete equations with a weak delay.* In *Dynamical System Modelling and Stability Investigation*. Kyjev, Ukrajina: University of Kyiv, UA, 2011. pp. 26-27. ISBN: 9667652009.

[4] DIBLÍK J., HALFAROVÁ H.: *Discrete systems of linear equations with weak delay.* In *XXIX International Colloquium on the Management of Educational Process*. Brno: 2011. pp. 1-4. ISBN: 978-80-7231-779- 0.

[5] DIBLÍK J., HALFAROVÁ H.: *Explicit general solution of planar linear discrete systems with constant coefficients and weak delays.* Adv. Difference Equ. 2013, Art. number: 50, doi:10.1186/1687-1847-2013-50, 1–29.

[6] DIBLÍK J., HALFAROVÁ H.: *General explicit solution of planar weakly delayed linear discrete systems and pasting its solutions.* Abstr. Appl. Anal. 2014, doi:10.1155/2014/627295, 1–37.

INTERVAL STABILITY OF NONLINEAR CONTROL SYSTEMS WITH AFTEREFFECT

Andriy Shatyrko
Taras Shevchenko National University of Kyiv,
Cybernetics faculty, Department of complex systems modelling
Volodymyrska str., 64, Kyiv, Ukraine, 01601
shatyrko.a@gmail.com

Abstract: *Sufficient conditions of interval absolute stability of nonlinear control systems described in terms of systems of the ordinary differential equations with delay argument, and also neutral type are obtained. The Lyapunov-Krasovskii functional method in the form of the sum of a quadratic component and integrals from nonlinearity is used at construction of statements.*

Keywords: stability, Lyapunov's method, deviating argument, nonlinear control systems.

INTRODUCTION

The actuality of absolute interval stability problem of the dynamical systems, mentioned in the present paper, proves to be true as a lot of interesting reports at the international congresses and conferences, and set of foreign publications, for example [1-6].

Problems of research of dynamical systems with it is inexact in the set parameters, or moreover, with vectors of speeds (the right-hand side of systems of the differential equations), accepting the values from some sets, interested researchers for a long time. Classical (Lyapunov) stability means investigation of solutions at indignations by the initial data [7]. Its various generalizations (uniform on time and phase variables, by parts variables, asymptotical, exponential, orbital etc.) also meant the unequivocal set of the law of dynamics of systems.

The solution of practical problems of control theory has caused occurrence concept "robust" (or interval) stability. Originally under robust stability it was understood asymptotical stability of the linear stationary differential equations of the higher order, under condition of a finding of their coefficients in the set intervals some beforehand. Interesting fundamental necessary and sufficient conditions of interval stability of the linear differential equations with it is inexact in the set parameters have been obtained at papers of Kharitonov V.L. [8-11]. However, at distribution of the obtained results to the dynamical systems, on differences equations and systems of the equations, systems with aftereffect, have arisen essential difficulties.

The solution of control problems in linear systems leads to a finding of function (scalar function) $u(x)$, at which feedback system

$$\dot{x}(t) = Ax(t) + bu(x(t))$$

should be asymptotical stable. Often this function depends on one scalar argument representing a linear combination of phase co-ordinates, and some scalar function from the first and third squares of a plane. Investigations of asymptotical stability of the systems

$$u(x(t)) = f(\sigma(t)), \quad \sigma(t) = c^T x(t),$$

i.e. systems

$$\dot{x}(t) = Ax(t) + bf(\sigma(t)), \ \sigma(t) = c^T x(t), \ t \geq 0.$$

with function $f(\sigma)$, lying in the set sector, became known as the absolute stability investigations of regulating (or control) systems.

Problems of control systems absolute stability have arisen in the middle of last century and are connected with problems of stabilization of programmed control at the set structure of control function [12,13,6]. The results giving absolute stability conditions, i.e. stability as a whole the zero solution for the set class of nonlinearity have been obtained in two directions.

One approach of investigations here is, so-called "frequency method", had development in Yakubovich V.A., Gelig A.H., Leonov G.A. works [14-17]. At the heart of a method is a study of behavior of some curve ("godograph") lies in complex area.

Other, alternative approach which has had development in works by Barbashin E.A., Martynyuk A.A., and other, is the Lyapunov second (direct) method with function type of "quadratic form plus integral from nonlinearity" [18-21].

Distribution of this method on systems with delay and neutral type has obtained in Khusainov D.Ya. and Shatyrko A.V. works [22-25]. Sufficient conditions of absolute interval stability have been constructed. At their construction the finite-dimensional method of Lyapunov's functions with a condition of Razumikhin B.S. [26] was used. The condition of Razumikhin B.S. facilitates construction of Lyapunov function. By means of this approach it is possible to estimate influence of aftereffect, i.e. to obtain the conditions of absolute interval stability depending from delay. However, the conditions of Razumikhin B.S. imposes rigid enough restrictions on aftereffect. And their use not always is effective.

At this paper we will use an alternative method of Lyapunov-Krasovskii functionals [6,11,27,28]. As the functionals the most effective are the integrated additives of a quadratic type. At this approach the obtained estimations become simpler. However, here as a point of phase space all piece of a trajectory is considered, therefore the approach does not allow to estimate influence of delay on absolute stability. Besides, the total derivative represents the quadratic form from phase co-ordinate and its prehistory. Therefore the matrix of the quadratic form of a total derivative has twice the big dimension.

1. DIRECT CONTROL SYSTEMS WITH TIME-DELAY ARGUMENT

At this section we will consider the system of direct control described by the differential equations with interval coefficients and with delay argument of next type

$$\begin{cases} \dot{x}(t) = (A + \Delta A)x(t) + (B + \Delta B)x(t - \tau) + bf(\sigma(t)) \\ \sigma(t) = c^T x(t) \end{cases} \tag{1}$$

Elements of matrices ΔA and ΔB also accept values from the fixed intervals

$$\begin{aligned} \Delta A &= \{\Delta a_{ij}\}, \ |\Delta a_{ij}| \leq \alpha_{ij}, \quad i, j = \overline{1, n}, \\ \Delta B &= \{\Delta b_{ij}\}, \ |\Delta b_{ij}| \leq \beta_{ij}, \quad i, j = \overline{1, n}. \end{aligned} \tag{2}$$

Nonlinear function $f(\sigma)$ satisfies to a "sector condition"

$$0 \le f(\sigma)\sigma \le k\sigma^2. \qquad (3)$$

Definition. The system (1) is called absolutely interval stable if it is absolutely stable for arbitrary matrices ΔA and ΔB from intervals (2).

Under absolute system stability we understand absolute stability of it trivial solution in sense of classical definitions [12,13].

At Khusainov D.Ya. and Shatyrko A.V. earlier papers conditions of interval stability of systems (1) with using of finite-dimensional Lyapunov's functions

$$V(x) = x^T Hx + \beta \int_0^{\sigma(x)} f(\xi)d\xi, \quad \sigma(x) = c^T x$$

have been obtained [22-25].

At the present paper we will construct conditions of interval stability of system (1) with the help of Lyapunov-Krasovskii functional

$$V[x(t)] = x^T(t)Hx(t) + \int_{-\tau}^{0} x^T(t+s)Gx(t+s)ds + \beta \int_0^{\sigma(t)} f(\sigma)d\sigma, \quad \sigma(t) = c^T x(t). \qquad (4)$$

We will use the following notations:

$\lambda_{\min}(\cdot)$, $\lambda_{\max}(\cdot)$ - accordingly the minimum and maximum own numbers of a matrix,

$\|\cdot\|$ -the Euclidean norm, $\|x(t)\|_2 = \left\{ \int_{-\tau}^{0} |x(t+s)|^2 ds \right\}$; $\|\Delta A\| = \max_{\Delta a_{ij}} \{|\Delta A|\}$; $\|\Delta B\| = \max_{\Delta b_{ij}} \{|\Delta B|\}$,

θ - zero-vector; Θ - zero-matrix.

Let's preliminary consider system with delay without "interval perturbations"

$$\begin{cases} \dot{x}(t) = Ax(t) + Bx(t-\tau) + bf(\sigma(t)) \\ \sigma(t) = c^T x(t) \end{cases} \qquad (5)$$

Theorem 1. Let is exists the positive definite matrices G, H and parameter $\beta > 0$ at which the matrix

$$S[G,H,\beta] = \begin{bmatrix} -A^T H - HA - G & -HB & -[Hb + \frac{1}{2}(\beta A^T + I)c] \\ -B^T H & G & \theta \\ -[Hb + \frac{1}{2}(\beta A^T + I)c]^T & \theta^T & \frac{1}{k} - \beta b^T c \end{bmatrix} \qquad (6)$$

is positive definite too. Then the system with delay without interval perturbations is absolutely stable.

Proof. As function $f(\sigma)$ satisfies to a condition (3), then for functional (4) following bilateral estimations are true

$$\lambda_{\min}(H)|x(t)|^2 + \lambda_{\min}(G)\|x(t)\|_2^2 \le V[x(t)] \le \left[\lambda_{\max}(H) + k\beta|c|^2\right]|x(t)|^2 + \lambda_{\max}(G)\|x(t)\|_2^2. \qquad (7)$$

We will calculate a total derivative of functional along system solutions. We will obtain

$$\frac{d}{dt}V[x(t)] = [Ax(t) + Bx(t-\tau) + bf(\sigma(t))]^T Hx(t) +$$
$$+ x^T(t)H[Ax(t) + Bx(t-\tau) + bf(\sigma(t))] + x^T(t)Gx(t) - x^T(t-\tau)Hx(t-\tau) +$$
$$\beta f(\sigma(t))c^T[Ax(t) + Bx(t-\tau) + bf(\sigma(t))].$$

Or, using so-called S-procedure [16],

$$\frac{d}{dt}V[x(t)] \le -\left(x^T(t), x^T(t-\tau), f(\sigma(t))\right)S[G,H,\beta]\left(x^T(t), x^T(t-\tau), f(\sigma(t))\right)^T,$$

where

$$S[G,H,\beta] = \begin{bmatrix} -A^TH - HA - G & -HB & -[Hb + \frac{1}{2}(\beta A^T + I)c] \\ -B^TH & G & \theta \\ -[Hb + \frac{1}{2}(\beta A^T + I)c]^T & \theta^T & \frac{1}{k} - \beta b^T c \end{bmatrix}.$$

If matrix $S[G,H,\beta]$ is positive definite, than

$$\frac{d}{dt}V[x(t)] \le -\lambda_{\min}(S[G,H,\beta])\left(|x(t)|^2 + |x(t-\tau)|^2 + |f(\sigma(t))|^2\right).$$

Thus, on the basis of Krasovskii weak theorem [28] if there are positive definite matrices G, H and $S[G,H,\beta]$, at which

$$\lambda_{\min}(H)|x(t)|^2 \le V[x(t)] \le \left[\lambda_{\max}(H) + k\beta|c|^2\right]|x(t)|^2 + \lambda_{\max}(G)\|x(t)\|_2^2.$$

$$\frac{d}{dt}V[x(t)] \le -\lambda_{\min}(S[G,H,\beta])|x(t)|^2,$$

then the system with delay (5) is absolutely stable.

Further we will obtain conditions of absolute interval stability of system (1).

Theorem 2. Let are exists the positive definite matrices G, H and parameter $\beta > 0$, at which the inequality is true

$$\lambda_{\min}(S[G,H,\beta]) > \|\Delta A\|H| + \sqrt{\|\Delta A\|^2|H|^2 + \|\Delta B\|^2|H|^2 + \frac{1}{4}\beta^2\|\Delta A\|^2|c|^2} \qquad (8)$$

Then the system (1) is absolutely interval stable.

Proof. As appears from a type of functional (4), for it bilateral estimations (7) are fair. We will calculate a total derivative of functional along solutions of system with "interval perturbations". We obtain

$$\frac{d}{dt}V[x(t)] = -\left(x^T(t), x^T(t-\tau), f(\sigma(t))\right)S[G,H,\beta]\left(x^T(t), x^T(t-\tau), f(\sigma(t))\right)^T +$$
$$+\left(x^T(t), x^T(t-\tau), f(\sigma(t))\right)\Delta S[G,H,\beta]\left(x^T(t)x^T(t-\tau), f(\sigma(t))\right)^T,$$

where

$$\Delta S[G,H,\beta] = \begin{bmatrix} \Delta A^TH + H\Delta A & H\Delta B & \frac{1}{2}\beta\Delta A^T c \\ \Delta B^TH & \Theta & \theta \\ \frac{1}{2}\beta c^T\Delta A & \theta^T & 0 \end{bmatrix}.$$

If matrix $S[G,H,\beta]$ is positive definite,

$$\frac{d}{dt}V[x(t)] \le -\lambda_{\min}\left(S[G,H,\beta]\right)\left(|x(t)|^2 + |x(t-\tau)|^2 + |f(\sigma(t))|^2\right) +$$
$$+ 2\|\Delta A\|\|H\|x(t)|^2 + 2\|\Delta B\|\|H\|x(t)\|x(t-\tau)| + \beta\|\Delta A\|\|c\|x(t)\|f(\sigma(t))|.$$

From here we have

$$\frac{d}{dt}V[x(t)] \le -\left[\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta A\|\|H|\right]x(t)|^2 -$$
$$- \lambda_{\min}\left(S[G,H,\beta]\right)|x(t-\tau)|^2 - \lambda_{\min}\left(S[G,H,\beta]\right)|f(\sigma(t))|^2 +$$
$$+ 2\|\Delta B\|\|H\|x(t)\|x(t-\tau)| + \beta\|\Delta A\|\|c|x(t)\|f(\sigma(t))|.$$

Let's break the first composed on two one and we will present the right part of an inequality in the form of the next sum

$$\frac{d}{dt}V[x(t)] \le -\left\{\alpha\left[\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta A\|\|H|\right]x(t)|^2 -\right.$$
$$\left. 2\|\Delta B\|\|H\|x(t)\|x(t-\tau)| + \lambda_{\min}\left(S[G,H,\beta]\right)|x(t-\tau)|^2\right\} -$$
$$- \left\{(1-\alpha)\left[\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta A\|\|H|\right]x(t)|^2 - \beta\|\Delta A\|\|c|x(t)\|f(\sigma(t))| +\right.$$
$$\left. + \lambda_{\min}\left(S[G,H,\beta]\right)|f(\sigma(t))|^2\right\},$$

where $0 < \alpha < 1$- some constant. Then, as appears from Sylvester's criterion, performance of inequalities will be a condition of absolute interval stability of system with delay

$$\begin{aligned}
&\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta A\|\|H| > 0,\\
&\alpha\left[\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta A\|\|H|\right]\lambda_{\min}\left(S[G,H,\beta]\right) - \left(\|\Delta B\|\|H|\right)^2 > 0, \qquad (9)\\
&(1-\alpha)\left[\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta A\|\|H|\right]\lambda_{\min}\left(S[G,H,\beta]\right) - \frac{1}{4}\beta\|\Delta A\|\|c| > 0
\end{aligned}$$

Let the ΔA such that the first inequality is executed. We will copy the second and third inequalities in a type

$$\alpha > \frac{\left(\|\Delta B\|\|H|\right)^2}{\left[\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta A\|\|H|\right]\lambda_{\min}\left(S[G,H,\beta]\right)}$$
$$\alpha < 1 - \frac{\frac{1}{4}\left(\beta\|\Delta A\|\|c|\right)^2}{\left[\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta A\|\|H|\right]\lambda_{\min}\left(S[G,H,\beta]\right)}.$$

And, if the inequality is true

$$\frac{\left(\|\Delta B\|\|H\|\right)^2}{\left[\lambda_{\min}\left(S[G,H,\beta]\right)-2\|\Delta A\|\|H\|\right]\lambda_{\min}\left(S[G,H,\beta]\right)}<$$

$$<1-\frac{\frac{1}{4}\left(\beta\|\Delta A\|\|c\|\right)^2}{\left[\lambda_{\min}\left(S[G,H,\beta]\right)-2\|\Delta A\|\|H\|\right]\lambda_{\min}\left(S[G,H,\beta]\right)},$$

than always exists $0<\alpha<1,$ at which the second and third inequalities (9) are true. And last inequality is equivalent to the following

$$\left(\|\Delta A\|\|H\|\right)^2+\frac{1}{4}\beta\|\Delta A\|\|c\|<\left[\lambda_{\min}\left(S[G,H,\beta]\right)-2\|\Delta A\|\|H\|\right]\lambda_{\min}\left(S[G,H,\beta]\right).$$

Let's copy it in a type

$$\left[\lambda_{\min}\left(S[G,H,\beta]\right)\right]^2-2\|\Delta A\|\|H\|\left[\lambda_{\min}\left(S[G,H,\beta]\right)\right]-$$

$$-\left\{\left(\|\Delta B\|\|H\|\right)^2+\frac{1}{4}\left(\beta\|\Delta A\|\|c\|\right)^2\right\}>0$$

It will be always true, if

$$\lambda_{\min}\left(S[G,H,\beta]\right)>\|\Delta A\|\|H\|+\sqrt{\|\Delta A\|^2|H|^2+\|\Delta B\|^2|H|^2+\beta^2\|\Delta A\|^2|c|^2}.$$

As from performance of last inequality performance of the first inequality (9) it is similar to the theorem 1, we obtain the statement (8) of theorem 2.

2. DIRECT CONTROL SYSTEMS OF NEUTRAL TYPE

We will consider the direct control system described by the differential equations with deviating argument of neutral type and with interval given coefficients of linear part

$$\begin{aligned}&\frac{d}{dt}\left(x(t)-Dx(t-\tau)\right)=\left(A+\Delta A\right)x(t)+\left(B+\Delta B\right)x(t-\tau)+bf\left(\sigma(t)\right)\\&\sigma(t)=c^T x(t).\end{aligned}\qquad(10)$$

Here a matrix D satisfies to a condition "difference operator stability", i.e. $|D|<1$, matrices ΔA and ΔB also can accept the values from the fixed intervals (2). Nonlinear scalar function of one argument $f(\sigma)$ lies in the set sector of the first and third quarter of coordinates plane (3).

In the present section for construction of absolute interval stability conditions we will use the functional of Lyapunov-Krasovskii of a following type

$$V[x(t)]=\left(x(t)-Dx(t-\tau)\right)^T H\left(x(t)-Dx(t-\tau)\right)+\int_{-\tau}^{0}x^T(t+s)Gx(t+s)ds+\beta\int_{0}^{\sigma(t)}f(\xi)d\xi,\qquad(11)$$

Let's preliminary consider system without interval perturbations

$$\begin{cases}\dfrac{d}{dt}\left(x(t)-Dx(t-\tau)\right)=Ax(t)+Bx(t-\tau)+bf\left(\sigma(t)\right)\\ \dot{\sigma}(t)=c^T x(t)\end{cases}\qquad(12)$$

Also we will obtain absolute stability conditions of system (12).

Let's denote

$$M[H]=\begin{bmatrix} H & HD \\ D^T H & D^T HD \end{bmatrix},$$

$$S[G,H,\beta]=\begin{bmatrix} -A^T H-HA-G & -HB+A^T HD & -\left[Hb+\frac{1}{2}\left(\beta A^T+I\right)c\right] \\ -B^T H+D^T HA & B^T HD+D^T HB+G & \theta \\ -\left[Hb+\frac{1}{2}\left(\beta A^T+I\right)c\right]^T & \theta^T & \frac{1}{k}-\beta b^T c \end{bmatrix} \quad (13)$$

Theorem 3. Let there exists positive definite matrices G, H, and parameter $\beta>0$, at which the matrix $S[G,H,\beta]$ also is positive definite. Then the system without interval perturbations (12) is absolutely stable in the metrics $\|x(t)\|_2$.

Proof. For Lyapunov-Krasovskii fnctional (11) following bilateral estimations are true

$$\lambda_{\min}(G)\|x(t)\|_2^2 \le V[x(t)] \le \lambda_{\max}(M[H])\left(|x(t)|^2+|x(t-\tau)|^2\right)+\lambda_{\max}(G)\|x(t)\|_2^2+\beta k|\sigma(t)|^2. \quad (14)$$

or

$$\lambda_{\min}(G)\|x(t)\|_2^2 \le V[x(t)] \le \left[\lambda_{\max}(M[H])+\beta k|c|^2\right]|x(t)|^2+\lambda_{\max}(M[H])|x(t-\tau)|^2+$$
$$+\lambda_{\max}(G)\|x(t)\|_2^2.$$

We will calculate a total derivative of functional (11) owing to system without interval perturbations. We obtain the following

$$\frac{d}{dt}V[x(t)]=[Ax(t)+Bx(t-\tau)+bf(\sigma(t))]^T H(x(t)+Dx(t-\tau))+$$
$$+(x(t)-Dx(t-\tau))^T H[Ax(t)+Bx(t-\tau)+bf(\sigma(t))]+x^T(t)Gx(t)-x^T(t-\tau)Gx(t-\tau)+$$
$$+\beta f(\sigma(t))c^T[Ax(t)+Bx(t-\tau)+bf(\sigma(t))].$$

Or, using S-procedure [16],

$$\frac{d}{dt}V[x(t)] \le -\left(x^T(t),x^T(t-\tau),f(\sigma(t))\right)S[G,H,\beta]\left(x^T(t),x^T(t-\tau),f(\sigma(t))\right)^T,$$

where matrix $S[G,H,\beta]$ is defined in (13). If it is positive definite, then

$$\frac{d}{dt}V[x(t)] \le -\lambda_{\min}(S[G,H,\beta])\left(|x(t)|^2+|x(t-\tau)|^2+|f(\sigma(t))|^2\right).$$

Thus we have system of inequalities

$$\lambda_{\min}(G)\|x(t)\|_2^2 \le V[x(t)] \le \left[\lambda_{\max}(M[H])+\beta k|c|^2\right]|x(t)|^2+\lambda_{\max}(G)\|x(t)\|_2^2.$$
$$\frac{d}{dt}V[x(t)] \le -\lambda_{\min}(S[G,H,\beta])|x(t)|^2.$$

And, on the basis of Krasovskii weak theorem [28], if there are positive definite matrices G, H, at which matrix $S[G,H,\beta]$ also it is positive definite, the system is absolutely stable in the metrics $\|x(t)\|_2$.

Further we will obtain absolute interval stability conditions of system (10).

Theorem 4. Let there exists positive definite matrices G, H and parameter $\beta > 0$, at which the next inequality is true

$$\lambda_{\min}\left(S[G,H,\beta]\right) > \left(\|\Delta A\|\,|H| + \|\Delta B\|\,|H|\right) + \\ + \sqrt{\left(\|\Delta A\|\,|H| + \|\Delta B\|\,|HD|\right)^2 + \left(\|\Delta B\|\,|H| + \|\Delta A\|\,|HD|\right)^2}\,. \tag{15}$$

Then the system (10) is interval absolutely stable in the metrics $\|x(t)\|_2$.

Proof. As appears from a type of functional (11), for it bilateral estimations (14) are true. We will calculate a total derivative of functional along solution of system with "interval perturbations". We obtain

$$\frac{d}{dt}V[x(t)] \le -\left(x^T(t), x^T(t-\tau), f(\sigma(t))\right)S[G,H,\beta]\left(x^T(t), x^T(t-\tau), f(\sigma(t))\right)^T + \\ + \left(x^T(t), x^T(t-\tau), f(\sigma(t))\right)\Delta S[G,H]\left(x^T(t)x^T(t-\tau), f(\sigma(t))\right)^T,$$

where

$$\Delta S[G,H] = \begin{bmatrix} \Delta A^T H + H\Delta A & -H\Delta B + \Delta A H D & \theta \\ -\Delta B^T H + D^T H \Delta A & \Delta B^T H D + D^T H \Delta B & \theta \\ \theta^T & \theta^T & 0 \end{bmatrix}.$$

if $S[G,H,\beta]$ is positive definite, then

$$\frac{d}{dt}V[x(t)] \le -\lambda_{\min}\left(S[G,H,\beta]\right)\left(|x(t)|^2 + |x(t-\tau)|^2 + |f(\sigma(t))|^2\right) + \\ 2\|\Delta A\|\,|H|\,|x(t)|^2 + 2\left(\|\Delta B\|\,|H| + \|\Delta A\|\,|HD|\right)|x(t)|\,|x(t-\tau)| + \|\Delta B\|\,|HD|\,|x(t-\tau)|^2.$$

From here we will have, that

$$\frac{d}{dt}V[x(t)] \le -\left[\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta A\|\,|H|\right]|x(t)|^2 + \\ 2\left(\|\Delta B\|\,|H| + \|\Delta A\|\,|HD|\right)|x(t)|\,|x(t-\tau)| - \left[\lambda_{\min}\left(S[G,H,\beta] - 2\|\Delta B\|\,|HD|\right)\right]|x(t-\tau)|^2 - \\ - \lambda_{\min}\left(S[G,H,\beta]\right)\left|f(\sigma(t))^2\right|.$$

Then, as appears from Sylvester's criterion [29], performance of system of inequalities will be a condition of absolute interval stability

$$\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta A\|\,|H| > 0, \\ \left[\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta A\|\,|H|\right]\left[\lambda_{\min}\left(S[G,H,\beta]\right) - 2\|\Delta B\|\,|HD|\right] - \\ - \left(\|\Delta B\|\,|H| + \|\Delta A\|\,|HD|\right)^2 > 0.$$

Let's copy the second inequality in a type

$$\lambda^2_{\min}\left(S[G,H,\beta]\right) - 2\left[\|\Delta A\|\,|H| + \|\Delta B\|\,|HD|\right]\lambda_{\min}\left(S[G,H,\beta]\right) - \left(\|\Delta B\|\,|H| - \|\Delta A\|\,|HD|\right)^2 > 0.$$

It will be true especially if there will be positive definite matrices G, H and parameter $\beta > 0,$ at which

$$\lambda_{\min}(S[G,H,\beta]) > [\|\Delta A\|H| + \|\Delta B\|HD|] +$$
$$+\sqrt{[\|\Delta A\|H| + \|\Delta B\|HD|]^2 + [\|\Delta B\|H| + \|\Delta A\|HD|]^2}.$$

From here the statement of the theorem 4 follows.

From the theorem 4 directly the consequence which is easier realized for check of conditions of interval stability follows.

Consequence. Let there exists positive definite matrices G, H and parameter $\beta > 0$, at which the inequality is true

$$\frac{\lambda_{\min}(S[G,H,\beta])}{\lambda_{\max}(H)} > (\|\Delta A\| + \|\Delta B\|) + \sqrt{(\|\Delta A\| + \|\Delta B\|D|)^2 + (\|\Delta B\| + \|\Delta A\|D|)^2}.$$

Then system (10) is absolute interval stable in the metrics $\|x(t)\|_2$.

CONCLUSION AND PROSPECTS

In the paper for the nonlinear systems of automatic control described in terms of the ordinary differential equations with delay and neutral type, and also having uncertainties in the set of linear parts, are received constructive algebraic criteria of interval absolute stability. At the expense of application of the alternative approach of Lyapunov-Krasovskii functional, form of estimations in sufficient conditions of interval stability are essentially simplified in comparison with obtained analogous one on the basis finite-dimensional Lyapunov's functions of Lur'e-Postnikov types [22,23,31,32].

In the chosen approach that results can be extended further on, a so-called, critical case (indirect control system) is perspective. Besides, applying the specified approach, similar results for the discrete systems are obtained [33,34]. It studying is actual enough recently. Also from the point of view of authors interest in the future represents construction of Lyapunov functions and Lyapunov-Krasovskii functionals, which are optimal in the classes by the set criteria of quality, for example [35].

It should also be noted next fact, if conditions of the Theorems 1-4 could not fulfill, it's not a dead-end situation. In such case, you can go for example to the solving the stabilization problem to a state of absolute stability [36,37].

All this confirms the viability and prospects of Lyapunov's direct method in the qualitative analysis of complex dynamical systems.

REFERENCES

[1]. LIU, D., MOLCHANOV, A. Criteria for robust absolute stability of time-varying nonlinear continuous-time systems. *Automatica.* 2002. V.38, №4. pp. 627-637.

[2]. SUN, J., DENG, F., LIY, Y. Robust absolute stability of general interval Lur'e type nonlinear control systems. *J. Syst. Engineering and Electronics*. 2001. V.12,№4. pp. 46-52.

[3]. YU, L., HAN, Q.L., YU, S., GAO, J. Delay-dependent conditions for robust absolute stability of uncertain time-delay systems. *Proc. IEEE Conf.Dec.Cont*. 2003. V.6. pp. 6033-6037.

[4]. GAO, J., PAN, H., DAI, W. A delay-dependent criterion for robust absolutely stability of uncertain Lurie type control systems. *Proc. World Congr. Intelligent Control And Automation (WCICA)*. 2004. V.1. pp. 928-930.

[5]. DONG, Y., LIU, J. Exponential stabilization of uncertain nonlinear time-delay systems. *Advances in Difference Equations*. doi:10.1186/1687-1847-2012:180.

[6]. LIBERZON, MR: Essays on the absolute stability theory. *Autom. Remote Control* 67(10), 1610-1644 (2006)

[7]. MALKIN, I.G. The theory of stability of movement. *M., Science*, 1966. 530 pp. (in Russian).

[8]. KHARITONOV, V.L. About asymptotic stability of position of balance of family systems of the linear differential equations. *J. Diff. equations*, №.11, 1978. pp. 2086-2088. (in Russian).

[9]. KHARITONOV, V.L., ZHABKO, A.P. Robust stability of time-delay systems. *IEEE Trans. On Automatic Control*, 39, pp. 2388-2397.

[10]. KHARITONOV, V.L. Criterion of stability of one family of quasipolynoms of delay type. *Automatics and telemechanics*, №2, 1991. pp. 73-82. (in Russian).

[11]. KHARITONOV, V.L. About preservation of property of global stability on shifts at the variation of parameters. *Annual Reviews in Control*, 23, *Published by Elsevier Science*1999. pp. 185-196.

[12]. LUR'E, A.I. Some problems in the theory of automatic control. *H.M. Stationary Office, London*, 1957.

[13]. AIZERMAN, M.A, GANTMAHER, F.R. Absolute stability of regulator systems. *Holden-Day, San Francisco*, 1964.

[14]. YAKUBOVICH, V.A., LEONOV, G.A, GELIG, A.Kh. Stability of stationary set ii control systems with discontinuous nonlinearities. *World Scientific, Singapore*, 2004.

[15]. POPOV, V.M. Hyperstability of control systems. Berlin, Springer-Verlag, 1973.

[16]. YAKUBOVICH, V.A. The S-procedure in the nonlinear control theory. *Bulletin of the Leningrad university. Mathematics-mechanics-astronomy*, №1, 1971. pp. 62-77. (in Russian).

[17]. BARKIN, A.I. Absolute stability of the control systems. *M., Book House «Librokom»*, 2012. 176 pp. (in Russian).

[18]. BARBASHIN E.A. Function of Lyapunov. *M., Science*, 1970. 240 pp. (in Russian).

[19]. KORENEVSKIY, D.G. *Stability of dynamic systems at random perturbation of parameters*. Algebraic criteria. K., Naukova Dumka, 1989. 208 pp. (in Russian).

[20]. MARTYNYUK, A.A., LAKSHMIKANTHAM, V., LEELA, S. Stability analysis of nonlinear systems. *Marcel Dekker, New York*, 1989.

[21]. MARTYNYUK, A.A. Stability by Liapunov's matrix function method with applications. *Marcel Dekker, New York*, 1998.

[22]. SHATYRKO, A.V., KHUSAINOV, D.Ya. Absolute interval stability of indirect regulating systems of neutral type. *Journal of automation and information science*. 2010. Vol.42, Iss.6, pp. 43-54.

[23]. SHATYRKO, A.V., KHUSAINOV, D.Ya. Investigation of Absolute Stability of Nonlinear Systems of Special Kind with Aftereffect by Lyapunov Functions Method. *Journal of automation and information science.* 2011. Vol.43, Iss.7, pp. 61-75.

[24]. KHUSAINOV, D.Ya., SHATYRKO, A.V. Lyapunov functions method in stability investigation of functional differential systems. *Kiev, Publ. by Kiev National University,* 1997. 236 pp. (in Russian).

[25]. SHATYRKO, A.V., KHUSAINOV, D.Ya. Stability of nonlinear control systems with aftereffect.*K., Inform. Analit. Agency Publ.,* 2012. 73 pp. ISBN 978-617-571-051-7 (in Ukrainian).

[26]. RAZUMIKHIN, B.S. Application of Lyapunov method to problems in stability of systems with delay. *Automatika i Telemehanika*, 21, 1960, pp. 740-749.(in Russian)

[27]. KRASOVSKII, N.N. On the applications of the second Lyapunov method for equations with delay. J. of Appl Math. Mech., 20, 1956, pp. 315-327.

[28]. EL'SGOL'TS, L.E., NORKIN S.B. Introduction to the theory of the differential equations with deviating argument. Academic Press, New York, 1973.

[29]. GANTMACHER,F.R. The Theory of Matrices (2 Volumes). (Matrix Theory, AMS Chelsea Publishing), 1998.

[30]. SHATYRKO, A.V. Absolute interval stability of neutral-type regulator systems. *Reports of the NASU* – 2011, №2. pp. 18–24. (in Russian).

[31]. SHATYRKO, A.V. Qualitative analyses of neutral type control systems under uncertainties from positions of Lyapunov functions. *Reports of the NASU* – 2012, №5. pp. 43–48. (in Russian).

[32]. SHATYRKO, A.V., KHUSAINOV, D.Ya. Absolute interval uniform stability investigations of neutral type systems by Lyapunov functions method. *Cybernetics and Computer Engineering* – 2011. Iss.166, pp. 3-14. (in Russian).

[33]. SHATYRKO, A.V. Absolute stability of discreet dynamical delay systems. *Ukrainian Math. Congress, 2009. Kyiv, Publ. of Inst. Math. NASU*.2010. pp. 203-213. ISBN 978-966-02-6192-1. (in Russian).

[34]. KHUSAINOV, D.Ya., SHATYRKO, A.V. Absolute stability conditions for difference systems. Bull. of Taras Shevchenko National University of Kyiv. Series Cybernetics. Iss.10, Kyiv, 2010. pp. 34 - 47. (in Ukrainian).

[35]. SHATYRKO A. Optimization method of absolute stability conditions constructing for nonlinear direct control systems. *Post-conference proceedings of selected papers extended version MITAV-2014. Brno, Czech Republic*, 2014. pp. 97-104. ISBN 978-80-7231-978-8

[36]. SHATYRKO, A., DIBLIK, J., KHUSAINOV, D., RUZICKOVA, M. Stabilization of Lur'e-type nonlinear control systems by Lyapunov-Krasovski functionals. *Advances in Difference Equations*. doi:10.1186/1687-1847-2012:229.

[37]. SHATYRKO, A., NOOIJEN, R.R.P., KOLECHKINA, A., KHUSAINOV, D. Stabilization of neutral-type indirect control systems to absolute stability state. *Advances in Difference Equations*. doi: 10.1186/s13662-015-0405-y.

RESEARCH OF STABILITY OF NEURAL NETWORK MODELS WITH DELAY BY THE SECOND LYAPUNOV METHOD

Sirenko A.S., Shakotjko T.I.
Faculty of Cybernetics, Taras Shevchenko National University of Kyiv,
Vladimirskayu Str., 64, Kyiv, 01601, Ukraine
sandrew@online.ua, trachuk_85@ukr.net

Abstract: *This report considered the dynamics of a neural network model that describes a system of differential equations and for study the stability using the method of Lyapunov functions with an additional condition Razumikhina. In rating the total derivative captured outside the diagonal elements.*

Keywords: differential equations, Lipschitz condition, neural network, asymptotically stable

INTRODUCTION

Mathematical models of the dynamics of neural networks described by nonlinear differential levels, with a dedicated asymptotic stable diagonal part reviewed in [1]. A more adequate model is system with delay.. It was designated in [2,3]. Apparatus of research such systems was chosen method of Lyapunov-Krasovskii functionals [2] and the method of comparison [3]. For research stability we using the method of Lyapunov functions with an additional condition Razumikhina [4,5]

1. MODEL OF THE PLANE. SYSTEM WITHOUT DELAY

We consider the following model of the dynamics of a neural network, described by a system of differential equations:

$$\dot{y}_1(t) = -a_{11}y_1(t) + f_{11}(y_1(t)) + f_{12}(y_2(t)) + b_1,$$
$$\dot{y}_2(t) = -a_{22}y_2(t) + f_{21}(y_1(t)) + f_{22}(y_2(t)) + b_2. \quad (1.1)$$

Where $a_{11} > 0$, $a_{22} > 0$ – constants, $f_{ij}(y)$, $i, j = \overline{1,2}$ - continuous functions, satisfy the condition Lipschitz

$$\left| f_{ij}(y + \Delta y) - f_{ij}(y) \right| \le L_{ij}|\Delta|, \; i, j = \overline{1,2}.$$

expected that the system of equations

$$-a_{11}y_1 + f_{11}(y_1) + f_{12}(y_2) + b_1 = 0, -a_{22}y_2 + f_{21}(y_1) + f_{22}(y_2) + b_2 = 0. \quad (1.2)$$

It has a unique solution point $M_0(y_1^0, y_2^0)$, $y_1^0 > 0$, $y_2^0 > 0$. After replacement $y_1(t) = x_1(t) + y_1^0$, $y_2(t) = x_2(t) + y_2^0$ we obtain:

$$\dot{x}_1(t) = -a_{11}x_1(t) + F_{11}(x_1(t)) + F_{12}(x_2(t)), \dot{x}_2(t) = -a_{22}x_2(t) + F_{21}(x_1(t)) + F_{22}(x_2(t)), \quad (1.3)$$

$$F_{11}(x_1(t)) = f_{11}(x_1(t) + y_1^0) - f_{11}(y_1^0), \; F_{12}(x_1(t)) = f_{12}(x_2(t) + y_2^0) - f_{12}(y_2^0),$$
$$F_{21}(x_1(t)) = f_{21}(x_1(t) + y_1^0) - f_{21}(y_1^0), \; F_{22}(x_1(t)) = f_{22}(x_2(t) + y_2^0) - f_{22}(y_2^0). \quad (1.4)$$

We have the following conditions for asymptotic stability.

Theorem 1.1. Let the system of equations (1.2) has a unique solution $M_0\left(y_1^0, y_2^0\right), y_1^0 > 0$, $y_2^0 > 0$ and exist constants $h_{11} > 0$, $h_{22} > 0$ in which the following conditions is performed

$$\left(a_{11} - L_{11}\right)h_{11} > 0,\ 4\left(a_{11} - L_{11}\right)\left(a_{22} - L_{22}\right)h_{11}h_{22} - \left(L_{12}h_{11} + L_{21}h_{22}\right)^2 > 0. \qquad (1.5)$$

Then the equilibrium state $M_0\left(y_1^0, y_2^0\right)$ is asymptotically stable.

Proof. For research stability of the equilibrium $M_0\left(y_1^0, y_2^0\right)$ use the quadratic Lyapunov function of the form $V\left(x_1, x_2\right) = h_{11}x_1^2 + h_{22}x_2^2$.

Its total derivative according to the system (1.4) has the form

$$\frac{d}{dt}V\left(x_1(t), x_2(t)\right) = 2h_{11}x_1(t)\left[-a_{11}x_1(t) + F_{11}\left(x_1(t)\right) + F_{12}\left(x_2(t)\right)\right] + 2h_{22}x_1(t)\left[-a_{22}x_1(t) + F_{21}\left(x_1(t)\right) + F_{22}\left(x_2(t)\right)\right].$$

Or $\frac{d}{dt}V\left(x_1(t), x_2(t)\right) = -2\left[a_{11}h_{11}x_1^2(t) + a_{22}h_{22}x_2^2(t)\right] + 2h_{11}x_1(t)\left\{F_{11}\left(x_1(t)\right) + F_{12}\left(x_2(t)\right)\right\} +$

$$+ 2h_{22}x_2(t)\left\{F_{21}\left(x_1(t)\right) + F_{22}\left(x_2(t)\right)\right\}.$$

Using the Lipschitz condition, we obtain

$$\frac{d}{dt}V\left(x_1(t), x_2(t)\right) \le -2\left[a_{11}h_{11}x_1^2(t) + a_{22}h_{22}x_2^2(t)\right] + 2h_{11}x_1(t)\left\{L_{11}\left|x_1(t)\right| + L_{12}\left|x_2(t)\right|\right\} +$$

$$+ 2h_{22}x_2(t)\left\{L_{21}\left|x_1(t)\right| + L_{22}\left|x_2(t)\right|\right\}.$$

We rewrite the expression whitch obtained in the form

$$\frac{d}{dt}V\left(x_1(t), x_2(t)\right) \le\le -2\left[\left(a_{11} - L_{11}\right)h_{11}x_1^2(t) - \left(L_{12}h_{11} + L_{21}h_{22}\right)\left|x_1(t)\right|\left|x_2(t)\right| + \left(a_{22} - L_{22}\right)h_{22}x_2^2(t)\right]$$

As the criterion of Sylvester [6], the condition of the total derivative is negative definite is implementation of inequalities

$$\left(a_{11} - L_{11}\right)h_{11} > 0,\ \left(a_{11} - L_{11}\right)\left(a_{22} - L_{22}\right)h_{11}h_{22} - \frac{1}{4}\left(L_{12}h_{11} + L_{21}h_{22}\right)^2 > 0,$$

i.e get the conditions (1.5).

2. MODEL IN THE PLANE. SYSTEM WITH DELAY

Let's consider the system on the plane with delay

$$\begin{aligned} \dot{y}_1(t) &= -a_{11}y_1(t) + f_{11}\left(y_1(t - \tau_{11})\right) + f_{12}\left(y_2(t - \tau_{12})\right) + b_1, \\ \dot{y}_2(t) &= -a_{22}y_2(t) + f_{21}\left(y_1(t - \tau_{21})\right) + f_{22}\left(y_2(t - \tau_{22})\right) + b_2. \end{aligned} \qquad (2.1)$$

We suppose that $\tau_{ij} > 0$, $i, j = 1,2$, $a_{11} > 0$, $a_{22} > 0$ and function $f_{ij}(y)$, $i, j = \overline{1,2}$ are continuous and satisfy a Lipschitz condition. Let make a replacement $y_1(t) = x_1(t) + y_1^0$, $y_2(t) = x_2(t) + y_2^0$ and the system (2.1) reduces to the form

$$\begin{aligned} \dot{x}_1(t) &= -a_{11}x_1(t) + F_{11}\left(x_1(t - \tau_{11})\right) + F_{12}\left(x_2(t - \tau_{12})\right), \\ \dot{x}_2(t) &= -a_{22}x_2(t) + F_{21}\left(x_1(t - \tau_{21})\right) + F_{22}\left(x_2(t - \tau_{22})\right). \end{aligned} \qquad (2.2)$$

And the research of the stability of the equilibrium position $M_0\left(y_1^0, y_2^0\right)$ has been reduced to the research of the stability of the zero equilibrium state of the system (2.2). We get the following conditions for asymptotic stability.

Theorem 2.1. Let the system of equations (2.2) has a unique solution $M_0\left(y_1^0, y_2^0\right)$ and there exist constants $h_{11} > 0$, $h_{22} > 0$ in which the following conditions performed

$$2\left[a_{11} - L_{11} - L_{12}\sqrt{\frac{h_{11}}{h_{22}}}\right]h_{11} > 0,\ 4\left[a_{11} - L_{11} - L_{12}\sqrt{\frac{h_{11}}{h_{22}}}\right]\left[a_{22} - L_{21}\sqrt{\frac{h_{22}}{h_{11}}} - L_{22}\right]h_{11}h_{22} -$$

$$-\left[\left(L_{11}\sqrt{\frac{h_{22}}{h_{11}}} + L_{12}\right)h_{11} + \left(L_{21} + L_{22}\sqrt{\frac{h_{11}}{h_{22}}}\right)h_{22}\right]^2 > 0. \qquad (2.3)$$

Then the equilibrium state $M_0\left(y_1^0, y_2^0\right)$ is asymptotically stable.

Proof. for the research of sustainability we will use the quadratic Lyapunov function $V(x_1, x_2) = h_{11}x_1^2 + h_{22}^2x_2^2$. For calculating the total derivative of the Lyapunov function by virtue of system (2.2) we will use B.S.Razumihina condition [4,5]. For the Lyapunov function $V(x_1, x_2) = h_{11}x_1^2 + h_{22}^2x_2^2$ it has the form

$$h_{11}x_1^2(s) + h_{22}^2x_2^2(s) = V(x_1(s), x_2(s)) < V(x_1(t), x_2(t)) = h_{11}x_1^2(t) + h_{22}x_2^2(t),\ s < t. \quad (2.4)$$

It follows that

$$|x_1(s)| < \sqrt{x_1^2(t) + \frac{h_{22}}{h_{11}}x_2^2(t)},\ |x_2(s)| < \sqrt{\frac{h_{11}}{h_{22}}x_1^2(t) + x_2^2(t)},\ s < t. \qquad (2.5)$$

The total derivative of the Lyapunov function by virtue of system (2.2) has the form

$$\frac{d}{dt}V\left(x_1(t), x_2(t)\right) = 2h_{11}x_1(t)\{-a_{11}x_1(t) + F_{11}(x_1(t-\tau_{11})) + F_{12}(x_2(t-\tau_{12}))\} +$$

$$+ 2h_{22}x_2(t)\{-a_{22}x_2(t) + F_{21}(x_1(t-\tau_{21})) + F_{22}(x_2(t-\tau_{22}))\}.$$

Using the Lipschitz condition, we obtain

$$\frac{d}{dt}V(x_1(t), x_2(t)) \le -2\left[a_{11}h_{11}x_1^2(t) + a_{22}h_{22}x_2^2(t)\right] + 2h_{11}x_1(t)\{L_{11}|x_1(t-\tau_{11})| + L_{12}|x_2(t-\tau_{12})|\} +$$

$$+ 2h_{22}x_2(t)\{L_{21}|x_1(t-\tau_{21})| + L_{22}|x_2(t-\tau_{22})|\}.$$

When we open the brackets we will get

$$\frac{d}{dt}V\left(x_1(t), x_2(t)\right) \le -2\left[a_{11}h_{11}x_1^2(t) + a_{22}h_{22}x_2^2(t)\right] + 2h_{11}x_1(t)L_{11}|x_1(t-\tau_{11})| + 2h_{11}x_1(t)L_{12}|x_2(t-\tau_{12})| +$$

$$+ 2h_{22}x_2(t)L_{21}|x_1(t-\tau_{21})| + 2h_{22}x_2(t)L_{22}|x_2(t-\tau_{22})|.$$

It is known that for arbitrary $A > 0$ and $B > 0$ following inequality holds

$$\sqrt{A^2 + B^2} < A + B. \qquad (2.6)$$

Using the conditions BS Razumihina (2.5) and inequality (2.6), we obtain

$$|x_1(t-s)| < \sqrt{x_1^2(t) + \frac{h_{22}}{h_{11}}x_2^2(t)} \le |x_1(t)| + \sqrt{\frac{h_{22}}{h_{11}}}|x_2(t)|,$$

$$|x_2(t-s)| < \sqrt{\frac{h_{11}}{h_{22}}x_1^2(t) + x_2^2(t)} \le \sqrt{\frac{h_{11}}{h_{22}}}|x_1(t)| + |x_2(t)|. \qquad (2.7)$$

It follows that

$$\frac{d}{dt}V(x_1(t),x_2(t)) \le -2[a_{11}h_{11}x_1^2(t)+a_{22}h_{22}x_2^2(t)]+2L_{11}h_{11}|x_1(t)|\left[|x_1(t)|+\sqrt{\frac{h_{22}}{h_{11}}}|x_2(t)|\right]+$$

$$+2L_{12}h_{11}|x_1(t)|\left[\sqrt{\frac{h_{11}}{h_{22}}}|x_1(t)|+|x_2(t)|\right]+2L_{21}h_{21}|x_2(t)|\left[|x_1(t)|+\sqrt{\frac{h_{22}}{h_{11}}}|x_2(t)|\right]+2L_{22}h_{22}|x_2(t)|[\sqrt{\frac{h_{11}}{h_{22}}}|x_1(t)|+|x_2(t)|].$$

Or

$$\frac{d}{dt}V(x_1(t),x_2(t)) \le -2\left[a_{11}-L_{11}-L_{12}\sqrt{\frac{h_{11}}{h_{22}}}\right]h_{11}x_1^2(t)+$$

$$+2\left[\left(L_{11}\sqrt{\frac{h_{22}}{h_{11}}}+L_{12}\right)h_{11}+\left(L_{21}+L_{22}\sqrt{\frac{h_{11}}{h_{22}}}\right)h_{22}\right]|x_1(t)||x_2(t)|-2[a_{22}-L_{21}\sqrt{\frac{h_{22}}{h_{11}}}-L_{22}]h_{22}x_1^2(t).$$

And, as the criterion of Sylvester [6], the condition of asymptotic stability will be the implementation of the system of inequalities

$$2\left[a_{11}-L_{11}-L_{12}\sqrt{\frac{h_{11}}{h_{22}}}\right]h_{11}>0,4\left[a_{11}-L_{11}-L_{12}\sqrt{\frac{h_{11}}{h_{22}}}\right]\left[a_{22}-L_{21}\sqrt{\frac{h_{22}}{h_{11}}}-L_{22}\right]h_{11}h_{22}-$$

$$-\left[\left(L_{11}\sqrt{\frac{h_{22}}{h_{11}}}+L_{12}\right)h_{11}+\left(L_{21}+L_{22}\sqrt{\frac{h_{11}}{h_{22}}}\right)h_{22}\right]^2>0,$$

i.e we get implementation of conditions (2.3).

3. SYSTEMS IN N-DIMENSIONAL SPACE

The most common case is the system delay in the n-dimensional space. The system has the form

$$\dot{y}_i(t)=-a_{ii}y_i(t)+\sum_{j=1}^{n}f_{ij}(y_j(t-\tau_{ij}))+b_i. \tag{3.1}$$

Let's make change $y_i(t)=x_i(t)+y_1^0$ and the system (3.1) reduces to the system

$$\dot{x}_i(t)=-a_{ii}(x_i(t)+y_i^0)+\sum_{j=1}^{n}f_{ij}(x_{ij}(t-\tau_{ij})+y_j^0)+b_i. \tag{3.2}$$

We rewrite it as

$$\dot{x}_i(t)=-a_{ii}x_i(t)+\sum_{j=1}^{n}F_{ij}(x_{ij}(t-\tau_{ij})), \tag{3.3}$$

$$F_{ij}(x_{ij}(t-\tau_{ij}))=f_{ij}(x_{ij}(t-\tau_{ij})+y_j^0)-f_{ij}(y_j^0) \tag{3.4}$$

And the research of the stability of the equilibrium position $M_0(y_1^0,y_2^0,...,y_n^0)$ has been reduced to the research of the stability of the zero equilibrium state of the system (3.3). We introduce the following notation

$$L_1=\frac{L_{11}}{\sqrt{h_{11}}}+\frac{L_{12}}{\sqrt{h_{22}}}+...+\frac{L_{1n}}{\sqrt{h_{nn}}},\ L_2=\frac{L_{21}}{\sqrt{h_{11}}}+\frac{L_{22}}{\sqrt{h_{22}}}+...+\frac{L_{2n}}{\sqrt{h_{nn}}},\dots,L_n=\frac{L_{n1}}{\sqrt{h_{11}}}+\frac{L_{n2}}{\sqrt{h_{22}}}+...+\frac{L_{nn}}{\sqrt{h_{nn}}}, \tag{3.5}$$

$$C=\begin{bmatrix}c_{11} & c_{12} & \dots & c_{1n}\\ c_{12} & c_{22} & \dots & c_{2n}\\ . & . & \dots & .\\ c_{1n} & c_{2n} & \dots & c_{nn}\end{bmatrix}, \tag{3.6}$$

$$c_{11}=2h_{11}\left(a_{11}-L_1\sqrt{h_{11}}\right),\ c_{12}=-\sqrt{h_{11}h_{22}}\left[L_1\sqrt{h_{11}}+L_2\sqrt{h_{22}}\right],\ldots,c_{1n}=-\sqrt{h_{11}h_{nn}}\left[L_1\sqrt{h_{11}}+L_n\sqrt{h_{nn}}\right],$$
$$c_{22}=2h_{22}\left(a_{22}-L_2\sqrt{h_{22}}\right),\ c_{23}=-\sqrt{h_{22}h_{33}}\left[L_2\sqrt{h_{22}}+L_3\sqrt{h_{33}}\right],\ldots,c_{2n}=-\sqrt{h_{22}h_{nn}}\left[L_2\sqrt{h_{22}}+L_n\sqrt{h_{nn}}\right],\ldots,$$
$$c_{n-1,n-1}=2h_{n-1,n-1}\left(a_{n-1,n-1}-L_{n-1}\sqrt{h_{n-1,n-1}}\right),\ c_{n-1,n}=-\sqrt{h_{n-1,n-1}h_{n,n}}\left[L_{n-1}\sqrt{h_{n-1,n-1}}+L_n\sqrt{h_{n,n}}\right],\quad(3.7)$$
$$c_{n,n}=2h_{nn}\left(a_{n,n}-L_n\sqrt{h_{n,n}}\right).$$

There have been the following conditions of stability

Theorem 3.1. Let the system of equations (3.1) has a unique equilibrium position $M_0\left(y_1^0,y_2^0,...,y_n^0\right)$ and exist onstants $h_{11}>0$, $h_{22}>0,\ldots,$ h_{nn} in which the conditions are performed

$$\Delta_1=h_{11}>0,\ \Delta_2=\begin{vmatrix}c_{11} & c_{12}\\ c_{12} & c_{22}\end{vmatrix}>0,\cdots\ \Delta_n=\begin{vmatrix}c_{11} & c_{12} & \ldots & c_{1n}\\ c_{12} & c_{22} & \ldots & c_{2n}\\ . & . & \ldots & .\\ c_{1n} & c_{2n} & \ldots & c_{nn}\end{vmatrix}>0\cdot\quad(3.8)$$

Then the equilibrium state $M_0\left(y_1^0,y_2^0,...,y_n^0\right)$ is asymptotically stable

Proof. For research stability of the equilibrium $M_0\left(y_1^0,y_2^0,...,y_n^0\right)$ w will use the Lyapunov function

$$V\left(x_1,x_2,...,x_n\right)=\sum_{n=1}^{n}h_{ii}x_i^2\ .\quad(3.9)$$

For calculating the total derivative of the Lyapunov function by virtue of the system (3.9) we will use B.S.Razumihina condition [4,5]. For the Lyapunov function (3.9) it has the form

$$\sum_{i=1}^{n}h_{ii}x_{ii}^2(s)=V\left(x_1(s),x_2(s),...,x_n(s)\right)<V\left(x_1(t),x_2(t),...,x_n(t)\right)=\sum_{k=1}^{n}h_{kk}x_k^2(t),\ s<t\ .\quad(3.10)$$

It follows that

$$|x_i(s)|<\sqrt{\sum_{k=1}^{n}\frac{h_{kk}}{h_{ii}}x_k^2(t)}\le\sqrt{\frac{h_{11}}{h_{ii}}}|x_1(t)|+\sqrt{\frac{h_{22}}{h_{ii}}}|x_2(t)|+...+\sqrt{\frac{h_{nn}}{h_{ii}}}|x_n(t)|,\ s<t,\ i=\overline{1,n}\ .\quad(3.11)$$

The total derivative of the Lyapunov function (3.9) by virtue of the system (3.3) has the form

$$\frac{d}{dt}V\left(x_1(t),x_2(t),...,x_n(t)\right)=\sum_{i=1}^{n}2h_{ii}x_i(t)\left\{-a_{ii}x_i(t)+\sum_{j=1}^{n}F_{ij}\left(x_j\left(t-\tau_{ij}\right)\right)\right\}.$$

Using the Lipschitz condition, we obtain

$$\frac{d}{dt}V\left(x_1(t),x_2(t),...,x_n(t)\right)\le-2\sum_{i=1}^{n}a_{11}h_{ii}x_{ii}^2(t)+2\sum_{i=1}^{n}h_{ii}x_i(t)\sum_{j=1}^{n}L_{ij}\left|x_j\left(t-\tau_{ij}\right)\right|.$$

Let's consider the second summand. Using B.S.Razumihina conditions (3.11) and inequality (2.6), we obtain

$$S_2=2\sum_{i=1}^{n}h_{ii}x_i(t)[\sum_{j=1}^{n}L_{ij}\left|x_j\left(t-\tau_{ij}\right)\right|]=2h_{11}x_1(t)\left[L_{11}\left|x_1\left(t-\tau_{11}\right)\right|+L_{12}\left|x_2\left(t-\tau_{12}\right)\right|+...+L_{1n}\left|x\left(t-\tau_{1n}\right)\right|\right]+$$
$$+2h_{22}x_2(t)\left[L_{21}\left|x_1\left(t-\tau_{21}\right)\right|+L_{22}\left|x_2\left(t-\tau_{22}\right)\right|+...+L_{2n}\left|x\left(t-\tau_{2n}\right)\right|\right]+...$$
$$+2h_{nn}x_n(t)\left[L_{n1}\left|x_1\left(t-\tau_{n1}\right)\right|+L_{n2}\left|x_2\left(t-\tau_{n2}\right)\right|+...+L_{nn}\left|x\left(t-\tau_{nn}\right)\right|\right]=$$

$$= 2h_{11}x_1(t)\Bigg\{L_{11}\left[\sqrt{\frac{h_{11}}{h_{11}}}|x_1(t)| + \sqrt{\frac{h_{22}}{h_{11}}}|x_2(t)| + \ldots + \sqrt{\frac{h_{nn}}{h_{11}}}|x_n(t)|\right] +$$

$$+ L_{12}\left[\sqrt{\frac{h_{11}}{h_{22}}}|x_1(t)| + \sqrt{\frac{h_{22}}{h_{22}}}|x_2(t)| + \ldots + \sqrt{\frac{h_{nn}}{h_{22}}}|x_2(t)|\right] + \ldots$$

$$+ L_{1n}\left[\sqrt{\frac{h_{11}}{h_{nn}}}|x_1(t)| + \sqrt{\frac{h_{22}}{h_{nn}}}|x_2(t)| + \ldots + \sqrt{\frac{h_{nn}}{h_{nn}}}|x_n(t)|\right]\Bigg\} +$$

$$+ 2h_{22}x_2(t)\{L_{21}[\sqrt{\frac{h_{11}}{h_{11}}}|x_1(t)| + \sqrt{\frac{h_{22}}{h_{11}}}|x_2(t)| + \ldots + \sqrt{\frac{h_{nn}}{h_{11}}}|x_n(t)|] +$$

$$+ L_{22}\left[\sqrt{\frac{h_{11}}{h_{22}}}|x_1(t)| + \sqrt{\frac{h_{22}}{h_{22}}}|x_2(t)| + \ldots + \sqrt{\frac{h_{nn}}{h_{22}}}|x_2(t)|\right] + \ldots$$

$$+ L_{2n}\left[\sqrt{\frac{h_{11}}{h_{nn}}}|x_1(t)| + \sqrt{\frac{h_{22}}{h_{nn}}}|x_2(t)| + \ldots + \sqrt{\frac{h_{nn}}{h_{nn}}}|x_n(t)|\right]\Bigg\} + 2h_{nn}x_n(t)\{L_{n1}[\sqrt{\frac{h_{11}}{h_{11}}}|x_1(t)| + \sqrt{\frac{h_{22}}{h_{11}}}|x_2(t)| + \ldots + \sqrt{\frac{h_{nn}}{h_{11}}}|x_n(t)|] +$$

$$+ L_{n2}\left[\sqrt{\frac{h_{11}}{h_{22}}}|x_1(t)| + \sqrt{\frac{h_{22}}{h_{22}}}|x_2(t)| + \ldots + \sqrt{\frac{h_{nn}}{h_{22}}}|x_2(t)|\right] + \ldots + L_{nn}\left[\sqrt{\frac{h_{11}}{h_{nn}}}|x_1(t)| + \sqrt{\frac{h_{22}}{h_{nn}}}|x_2(t)| + \ldots + \sqrt{\frac{h_{nn}}{h_{nn}}}|x_n(t)|\right]\Bigg\}.$$

We introduce the following notation

$$L_1 = \frac{L_{11}}{\sqrt{h_{11}}} + \frac{L_{12}}{\sqrt{h_{22}}} + \ldots + \frac{L_{1n}}{\sqrt{h_{nn}}},\ L_2 = \frac{L_{21}}{\sqrt{h_{11}}} + \frac{L_{22}}{\sqrt{h_{22}}} + \ldots + \frac{L_{2n}}{\sqrt{h_{nn}}}, \ldots,\ L_n = \frac{L_{n1}}{\sqrt{h_{11}}} + \frac{L_{n2}}{\sqrt{h_{22}}} + \ldots + \frac{L_{nn}}{\sqrt{h_{nn}}}.$$

Then

$$S \le 2h_{11}x_1(t)L_1\left\{\sqrt{h_{11}}|x_1(t)| + \sqrt{h_{22}}|x_2(t)| + \ldots + \sqrt{h_{nn}}|x_n(t)|\right\} +$$

$$+ 2h_{22}x_2(t)L_2\left\{\sqrt{h_{11}}|x_1(t)| + \sqrt{h_{22}}|x_2(t)| + \ldots + \sqrt{h_{nn}}|x_n(t)|\right\} + \ldots$$

$$+ 2h_{nn}x_n(t)L_n\left\{\sqrt{h_{11}}|x_1(t)| + \sqrt{h_{22}}|x_2(t)| + \ldots + \sqrt{h_{nn}}|x_n(t)|\right\}.$$

Rearranging the quadratic terms, we obtain

$$S \le 2h_{11}L_1\sqrt{h_{11}}x_1^2(t) + 2\left[h_{11}L_1\sqrt{h_{22}} + h_{22}L_2\sqrt{h_{11}}\right]x_1(t)\|x_2(t)| +$$

$$+ 2\left[h_{11}L_1\sqrt{h_{33}} + h_{33}L_3\sqrt{h_{11}}\right]x_1(t)\|x_3(t)| + \ldots + 2\left[h_{11}L_1\sqrt{h_{nn}} + h_{nn}L_{nn}\sqrt{h_{11}}\right]x_1(t)\|x_n(t)| +$$

$$+ 2h_{22}L_2\sqrt{h_{22}}x_2^2(t) + 2\left[h_{22}L_2\sqrt{h_{33}} + h_{33}L_3\sqrt{h_{22}}\right]x_2(t)\|x_3(t)| + \ldots + 2\left[h_{22}L_2\sqrt{h_{nn}} + h_{nn}L_n\sqrt{h_{22}}\right]x_2(t)\|x_n(t)| + \ldots$$

$$+ 2h_{n-1,n-1}L_{n-1}\sqrt{h_{n-1,n-1}}x_{n-1}^2(t) + 2\left[h_{n-1,n-1}L_{n-1}\sqrt{h_{n,n}} + h_{n,n}L_n\sqrt{h_{n-1,n-2}}\right]x_{n-1}(t)\|x_n(t)| + + 2h_{n,n}L_n\sqrt{h_{n,n}}x_n^2(t).$$

We transform this expression the following way

$$S \le 2h_{11}L_1\sqrt{h_{11}}x_1^2(t) + 2\sqrt{h_{11}h_{22}}\left[L_1\sqrt{h_{11}} + L_2\sqrt{h_{22}}\right]x_1(t)\|x_2(t)| +$$

$$+ 2\sqrt{h_{11}h_{33}}\left[L_1\sqrt{h_{11}} + L_3\sqrt{h_{33}}\right]x_1(t)\|x_3(t)| + \ldots + 2\sqrt{h_{11}h_{nn}}\left[L_1\sqrt{h_{11}} + L_n\sqrt{h_{nn}}\right]x_1(t)\|x_n(t)| +$$

$$+ 2h_{22}L_2\sqrt{h_{22}}x_2^2(t) + 2\sqrt{h_{22}h_{33}}\left[L_2\sqrt{h_{22}} + L_3\sqrt{h_{33}}\right]x_2(t)\|x_3(t)| + \ldots + 2\sqrt{h_{22}h_{nn}}\left[L_2\sqrt{h_{22}} + L_{nn}\sqrt{h_{nn}}\right]x_2(t)\|x_n(t)| + \ldots$$

$$+ 2h_{n-1,n-1}L_{n-1}\sqrt{h_{n-1,n-1}}x_{n-1}^2(t) + 2\sqrt{h_{n-1,n-1}h_{nn}}\left[L_{n-1}\sqrt{h_{n,n}} + L_n\sqrt{h_{n-1,n-1}}\right]x_{n-1}(t)\|x_n(t)| + + 2h_{n,n}L_n\sqrt{h_{n,n}}x_n^2(t).$$

Let's introduce the following notation

$$C = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{12} & c_{22} & \dots & c_{2n} \\ . & . & \dots & . \\ c_{1n} & c_{2n} & \dots & c_{nn} \end{bmatrix},$$

$$c_{11} = 2h_{11}\left(a_{11} - L_1\sqrt{h_{11}}\right),\ c_{12} = -\sqrt{h_{11}h_{22}}\left[L_1\sqrt{h_{11}} + L_2\sqrt{h_{22}}\right],\dots, c_{1n} = -\sqrt{h_{11}h_{nn}}\left[L_1\sqrt{h_{11}} + L_n\sqrt{h_{nn}}\right],$$

$$c_{22} = 2h_{22}\left(a_{22} - L_2\sqrt{h_{22}}\right),\ c_{23} = -\sqrt{h_{22}h_{33}}\left[L_2\sqrt{h_{22}} + L_3\sqrt{h_{33}}\right],\dots, c_{2n} = -\sqrt{h_{22}h_{nn}}\left[L_2\sqrt{h_{22}} + L_n\sqrt{h_{nn}}\right],\dots,$$

$$c_{n-1,n-1} = 2h_{n-1,n-1}\left(a_{n-1,n-1} - L_{n-1}\sqrt{h_{n-1,n-1}}\right),\ c_{n-1,n} = -\sqrt{h_{n-1,n-1}h_{n,n}}\left[L_{n-1}\sqrt{h_{n-1,n-1}} + L_n\sqrt{h_{n,n}}\right],\ c_{n,n} = 2h_{n,n}\left(a_{n,n} - L_n\sqrt{h_{n,n}}\right).$$

Then, for a total derivative of the Lyapunov function due to a delay system (3.3) will have the inequality

$$\frac{d}{dt}V\left(x_1(t), x_2(t), \dots x_n(t)\right) \le -\left(|x_1(t)|, |x_2(t)|, \dots, |x_n(t)|\right)C\left(|x_1(t)|, |x_2(t)|, \dots, |x_n(t)|\right)^T .$$

And the condition of stability is a positive definite of the matrix *C*. As the Sylvester criterion, it is necessary and sufficient that all principal diagonal minors were positive, i.e, the conditions of Theorem 3.1. ned to be performed.

We can show that in the particular case of n-terms of the asymptotic stability of (3.8) coincide with the conditions of stability (2.3) in the plane.

REFERENCES

[1] S. HAYKIN Neural networks: a complete course, 2nd edition, M.: Publishing House "Williams", 2006. 1104 pp.

[2] GOPALSAMY K. *Leakage Delays in BAM*. Journal of Mathematical Analysis and Applications, 325 (2007), pp.1117-1132.

[3] BEREZANSKY L., IDELS L., TROIB L. *Global dynamics of the class on nonlinear nonautonomous systems with time-varying delays*. Nonlinear Anal. 74 (2011), No. 18, pp. 7499-7512.

[4] D.Ya. KHUSAINOV, SHATYRKO A.V. *Method of Lyapunov functions in stability analysis of functional-differential systems*. Kiev, Izd. of Kiev University, 1997. 236 pp.

[5] B.S. RAZUMIKHIN *Stability hereditarity systems.* M., Nauka, 1988. 112 pp.

[6] GANTMAKHER F.R. Matrix theory. M., Nauka, 1988. 548 pp.

APPLICATION OF NON-LINEAR PROGRAMMING TO OPTIMIZE TECHNOLOGICAL PROCESS

Alena VAGASKÁ[a], Miroslav GOMBÁR[b], Erika FECHOVÁ[a], Peter MICHAL[b]

[a] Department of Mathematics, Informatics and Cybernetics, Technical University of Košice, Faculty of Manufacturing Technologies with a seat in Prešov
Bayerova 1, 080 01 Prešov, Slovakia
alena.vagaska@tuke.sk, erika.fechova@tuke.sk

[b] Department of Mechanical Engineering, Institute of Technology and Business in České Budějovice,
Okružní 10, 370 01 České Budějovice, Czech Republic
gombar.mirek@gmail.com,
michal.peter.pm@gmail.com

Abstract: *The authors of the paper deal with the application of mathematical and statistical methods and Design of Experiments (DOE) in order to identify and analyse factors affecting the process of electrolytic alkaline zinc plating at a current density of 0.5 A.dm^{-2}. Based on the DOE methodology according to the central composite design, the set of experiments containing 40 runs has been performed. The influence of seven input factors on the final thickness of formed zinc layer has been investigated. In this paper, the mathematical-statistical model predicting the thickness of the formed layer is presented. In order to save time, as the possibility of increasing the efficiency of the technological process, non-linear programming was used to optimize the zincing process.*

Keywords: design of experiments (DOE), mathematical – statistical model, electrolytic alkaline zinc plating, significant factors, optimization.

INTRODUCTION

Technological processes of surface treatment belong to multifactorial systems, present complex non-linear processes actuating several technological, physical, chemical and material effects and their mutual interactions. That is why the analysis of these processes by classic methods appears to be non-efficient and many times leads to incorrect conclusions. At the process analysis, observation, examination, comparison and synthesis (optimization and forecasting), we find the basis in determination of bonds and relations between input and output parameters. It is significant to identify if certain factors (input parameters) have an influence on observed parameter (response). Then it is necessary to find such levels of factors in order to reach the optimum (maximum, minimum) of the observed parameter [1], [2]. To solve such practical problems it is more suitable to use experimental and statistic approach than determination approach. The analysis and synthesis in conditions of incomplete information is carried out at experimental and statistic approach of process analysis of surface adjustments. Even though the nature of examined process is not completely known, incomplete information for setting optimal conditions is updated by the experiment and known data. Wide use of various experimental methods is the consequence of incomplete

information and continual improvement of old and creation of new objects, processes and procedures [3].

The process of electrolytic alkaline zinc plating, where at the right choice of technological factors it is possible to create such a protective layer of material that will have requested thickness and properties and it will fulfil defined criteria (e.g. resistance against corrosion), also belongs to multifactorial and non-linear systems. Identification and analysis of the factors functioning in this process and observing their influence on created layer became the object of our scientific research and experimental work [4], [5], [6] and it is also solved in the paper. In contrast to majority of scientific and expert works on the problem [7], [8], [9], [10], where the selected parameter of created layer in relation to only one factor is analysed, the paper focuses on deeper and more complex identification of influences of several factors and their interactions functioning in the process of electrolytic alkaline zinc plating, which did not go without the use of Design of Experiments (DOE). In contrast to COST approach Design of Experiments (DOE) enables us to observe in given time common influence of several factors on the response and find optimal combination of setting the values of input process parameters. The change of only one selected factor in given time is considered at COST approach within the frame of experimental work (COST is an acronym of English expression "*c*onsider *o*ne *s*eparate factor at a *t*ime"), which is inefficient approach, because it does not provide necessary information in order to reach real optimum, experimental work is overpriced.

1. EXPERIMENTAL PART

S355J0 material was used to carry out the experiment. Within the frame of individual experiments above mentioned method of electrolytic alkaline zinc plating at current density of 0.5 $A.dm^{-2}$ was used. Zinc electrolyte was used, which is characteristic of its high depth efficiency, low zinc concentration and high coating ability. Zinc is currently an available option at the protection of metals from corrosion and creating special properties of material surface. Zinc anodes are placed into a separate dissolution bath, where it is possible to regulate the zinc content by sinking and lifting of anodes. Zinc is brought into the coating bath through filter by the circulation circuit. During alkaline coating it was necessary to secure the components of glitter additives in requested concentration in the electrolyte. Zink plating of samples was based on DOE with selected central composite plan with 40 individual experiments. We were interested in the influence of 7 input factors functioning on thickness of created zinc layer, i.e. functional dependency $\hat{y} = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$, where x_1 – is the amount of NaOH in the electrolyte, x_2 – is the amount of ZnO in the electrolyte, x_3 – the amount of glitter additive Pragogal Zn3401 in the electrolyte, x_4 – the amount of glitter additive Pragogal Zn3402 in the electrolyte, x_5 – electrolyte temperature, x_6 – time of zinc plating, x_7 – voltage. Experimental conditions and individual levels of individual variables (factors) can be found in Table 1. The thickness of layer coating using the digital thickness meter MINITEST 4000 was measured at individual samples in selected experimental points. Experimentally obtained data presented an input matrix for the further statistic processing.

Design of Experiments, which was used to identify significant factors influencing the thickness of created layer, enabled us to obtain maximum amount of information with high statistic and numeric precision at optimal number of individual experiments [2].

Coded factor	Factor	Unit	Factor level in planned experiment				
			-2,21	-1	0	1	2,21
x_1	$m(NaOH)$	$[g \cdot l^{-1}]$	19.33	80.00	120.00	180.00	240.67
x_2	$m(ZnO)$	$[g \cdot l^{-1}]$	3.15	8.00	12.00	16,00	20.85
x_3	$m(Zn\ 3401)$	$[ml \cdot l^{-1}]$	2.18	4.00	5.50	7.00	8.82
x_4	$m(Zn\ 3402)$	$[ml \cdot l^{-1}]$	0.68	2.50	4.00	5.50	7.32
x_5	T	[°C]	-0.13	12.00	22.00	32.00	44.13
x_6	t	[min]	1.15	6.00	10.00	14.00	18.85
x_7	U	[V]	0.79	2.00	3.00	4.00	5.21

Table 1. Indication and values of technological factors

Individual experiments were carried out on the basis of created matrix of experimental plan as a combination of individual levels of 7 input factors in accordance with Table 1, in which experimental conditions can be found. Individual experiments were carried out in random order. This randomisation is needed because of minimizing systematic errors or preventing subjective preferring of some of the input factor levels. Orthogonality of experimental plan was verified by means of the scalar products, i.e. all matrix columns of experimental plan must be perpendicular to each other and non-zero in order to avoid the wrong indication of statistic non-significance of regressors [6]. Well known transpose relation [6], due to which original physical units can be transposed to non-dimensional form, was used to norm (code) the basic factors.

2. RESULTS AND DISCUSSION

Exploring analysis, screening analysis, dispersion analysis and DoE analysis were carried out based on statistical analysis of experimentally obtained data. By using software products such as Matlab, Statistica, JMP or QC - Expert we recognized significant factors that have influence on the final thickness of AAO layer, analysed their interactions and obtained the shape and coefficients of mathematical and statistic models that predict the thickness of created layer at changing factor levels. Data analysis was carried out with statistically correct approach involving the analysis of basic conditions and following analysis of the classic regression triplet: data, model, residuals. That is why it can be said there was no numeric and statistic incorrectness of the results when deducing and interpreting the results, which was also confirmed by practical experiences in the area of surface treatment.

Basic analysis of obtained results of measuring thickness of created layer at individual experiments results from dispersion analysis (ANOVA), Table 2.

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	6	207.5383	34.5897	5.9572	0.0003*
Error	33	191.6105	5.8064		
C. Total	39	399.1488			

Table 2. ANOVA table for proposed prediction model

It can be judged from the table of dispersion analysis that variability caused by random errors is markedly lower than variability of measured values explained by the model and value of obtained significance level (Prob > F) points out adequacy of used model based on Fisher–Snedecor test. The further testing of used model by so called error test of insufficient model

adjustment, where we observe residual dispersion and dispersion of measured data inside the groups and test if regression model sufficiently describes observed dependency, can be found in Table 3. Considering the obtained value of significance 0.3486 by Fisher test it can be said that the model sufficiently describes variability of experimentally obtained data. Model significance is confirmed, dispersion of residual values is lower than dispersion inside individual groups at the selected significance level of $\alpha = 0.05$.

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F	Max RSq
Lack Of Fit	28	171.2105	6.11466	1.4987	0.3486	0.9489
Pure Error	5	20.4	4.08			
Total Error	33	191.6105				

Table 3. Table of error of insufficient model adjustment

The following Table 4 presents assessment of model parameters with testing of significance of individual effects and their combination at significance level $\alpha = 0.05$ based on above mentioned conditions and their completion (Table 2 and Table 3).

Term	Estimate	Std Error	t Ratio	Prob>|t|	Lower 95%	Upper 95%	VIF
Intercept	14.10792	0.599205	23.54	<.0001*	12.88882	15.32701	.
x_7	1.483657	0.432078	3.43	0.0016*	0.604588	2.362725	0.985741
x_6	1.043233	0.432078	2.41	0.0215*	0.164164	1.922301	0.985741
$x_1 . x_1$	-1.06999	0.368621	-2.9	0.0065*	-1.81995	-0.32002	0.911005
$x_1 . x_2$	-0.78191	0.512519	-1.53	0.1366	-1.82464	0.260817	0.985741
$x_5 . x_4$	0.647948	0.512519	1.26	0.215	-0.39478	1.690676	0.985741
$x_4 . x_4$	-0.79972	0.368621	-2.17	0.0373*	-1.54969	-0.04976	0.911005

Table 4. Table of assessment of model parameters
(x_7 – voltage, x_6 – time of zinc plating, x_1 – amount of NaOH, x_2 – amount of ZnO, x_4 – amount of glitter additive Pragogal Zn3402, x_5 – electrolyte temperature, *intercept* – absolute term of the model, * - factor or factor combination is significant at selected significance level of 5 %)

It is obvious from the table of assessment of model parameters that voltage input and time of zinc plating have the main influence on the thickness of created zinc layer. Except for individual functioning factors the second power of NaOH in the electrolyte and amount of glitter additive Pragogal Zn3402 have also significant impact. The absolute term of the model, which contains all "neglected" functioning factors in the process of electrolytic alkaline zinc plating, has the highest importance. *VIF* (Variance Inflation Factor) or inflation factors of dispersion of regressors are important indicators from the point of view of statistic criterion of non-orthogonality [7]. It is valid that *VIF* is lower or equal to 1 (predictors are not correlated, plan is uncorrelated and orthogonal), higher than 1 but lower than 5 (indication of medium correlation and plan non-orthogonality), higher than 5 but lower than 10 (significant correlation and plan non-orthogonality) and finally higher than 10 (multi correlation of regressors and plan non-orthogonality). If *VIF* >1, assessment of regression coefficients is numerically correct, but their *p – values*, which are defined as diagonal elements of inversion correlation matrix (j = 1...p), are not correct.

$$VIF_j = D_{jj} = diag(\mathbf{D}) = diag(\mathbf{R}^{-1}) \tag{1}$$

Since results point out uncorrelation of predictors and orthogonality of experimental plan, based on Table 4 prediction equation for the thickness of created layer ($\hat{y} = th$) in coded form can be expressed as

$$\begin{aligned} \hat{y} &= 14{,}10792 + 1{,}483657 \cdot x_7 + 1{,}043233 \cdot x_6 - 1{,}06999 \cdot x_1^2 \\ &- 0{,}78191\, x_1 \cdot x_2 + 0{,}647948\, x_5 \cdot x_4 - 0{,}79972 \cdot x_4^2 \end{aligned} \tag{2}$$

To set up the prediction relation in the natural scale it is necessary to realize that within the process of analysis used factors were coded by DoE norming in a coded scale:

$$x_d(i) = \frac{x(i) - \frac{x_{\max} + x_{\min}}{2}}{\frac{x_{\max} - x_{\min}}{2}} \tag{3}$$

where $x_d(i)$ is normed variable according to DoE, $x(i)$ - original basic variable, where $i = 1,2,3 \ldots n$, n - the number of basic factors, $x_{\max}$ - maximum value of original variable $x(i)$, $x_{\min}$ - minimum value of original valuable $x(i)$.

Considering transpose relation and statistic prediction equation (3) it is possible to express prediction relation describing observed dependency for current density of 0.5 A.dm^{-2} as

$$\begin{aligned} th &= m(Zn3402) \cdot 0{,}386 - 0{,}035 \cdot T + 0{,}671 \cdot U + 0{,}118 \cdot t + 0{,}034 \cdot m(NaOH) \\ &+ 0{,}112 \cdot m(ZnO) - 8{,}736 \cdot 10^{-5} \cdot \left(m(NaOH)\right)^2 - 0{,}073 \cdot \left(m(Zn3402)\right)^2 \\ &- 7{,}983 . 10^{-4} (m(NaOH) \cdot m(ZnO)) + 8{,}819 . 10^{-3} \cdot (m(Zn3402) \cdot T)) + 6{,}806 \end{aligned} \tag{4}$$

The prediction equation will serve as the basis to optimize the process by non-linear programming.

3. PROCESS OPTIMIZATION

The nature of optimization problems lies in determining such a combination of values of individual factors, at which "the best" value of optimization parameter is obtained. The obtained factor values are called optimal values. Optimization problems are highly significant at proposals of technological processes as well as projecting engineering objects, their realization and during their operation. In term of surface treatment of metals the time of the process duration is one of the most important parameters that determine the efficiency of the entire process. If we manage to minimize the time needed to create the layer with requested thickness at setting functioning factors, economic profit can be maximized at securing requested quality. If the time of zinc plating is expressed from equation (3), an optimization (criterion) function is obtained.

The basic task of non-linear optimization is to find the minimum of the problem defined as

$$\min_x f(x) \begin{cases} c(x) \le 0 \\ ceq(x) = 0 \\ \mathbf{A}.\mathbf{x} \le b \\ \mathbf{Aeq}.\mathbf{x} = \mathbf{beq} \\ \mathbf{lb} \le x \le \mathbf{ub} \end{cases} \tag{5}$$

where **x**, **beq**, **lb** and **ub** are vectors, **A** and **Aeq** are matrices, *c*(x) and *ceq*(x) are vector functions and *f(x)* is a scalar function. Functions *f*(x), *c*(x) and *ceq*(x) are non-linear functions [8]. Conditional inequations are obtained at defining of fringe conditions of the process of electrolytic alkaline zinc plating considering data presented in Table 1 and Table 4 and nature of the process

$$80 \le m(NaOH) \le 130 \tag{6}$$

$$7{,}5 \le m(ZnO) \le 18 \tag{7}$$

$$2 \le m(Pragogal\,Zn\ 3402) \le 6{,}5 \tag{8}$$

$$12 \le T \le 28 \tag{9}$$

$$2 \le U \le 5 \tag{10}$$

To solve optimization problem (4) non-linear programming in Matlab was used. In consideration of requested thickness of the layer of 12 μm as the most often requested thickness respecting fringe conditions (6) to (10), optimal time of 11.305 [min] is obtained at $m(NaOH)$=112.585 [$g \cdot l^{-1}$], $m(ZnO)$=18 [$g \cdot l^{-1}$], $m(Pragogal\ Zn\ 3402)$=3.392 [$ml \cdot l^{-1}$], T=12 [^oC] a U=5 [V]. The graphic output of optimization can be found in Fig. 1 - Fig. 5.

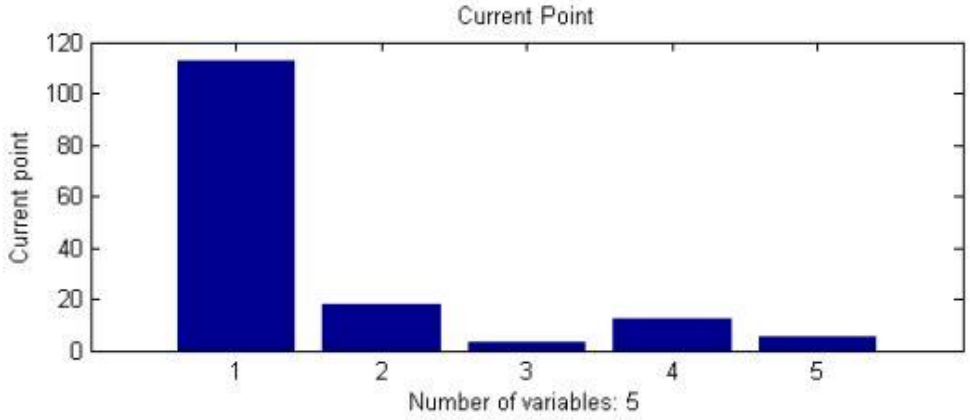


Fig. 1. Graph of current points of optimization of zinc plating time for layer thickness of 12 μm

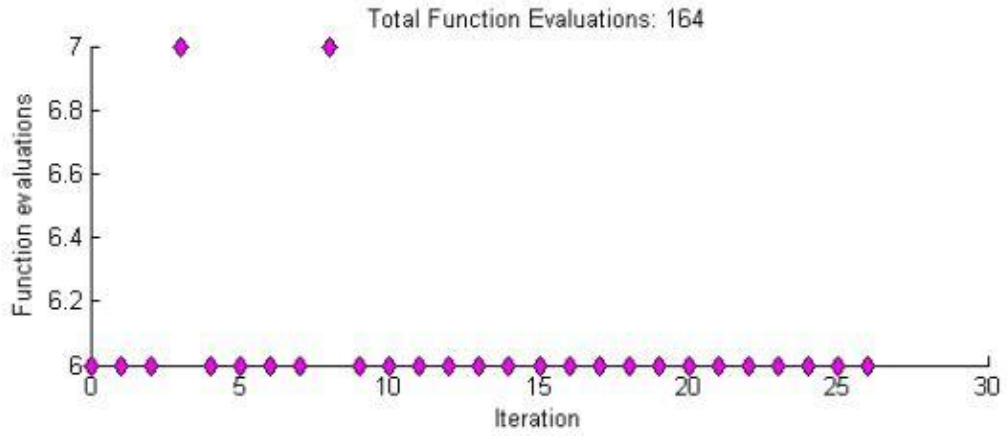


Fig. 2. Entire function evaluations of optimization of zinc plating time for layer thickness of 12 µm

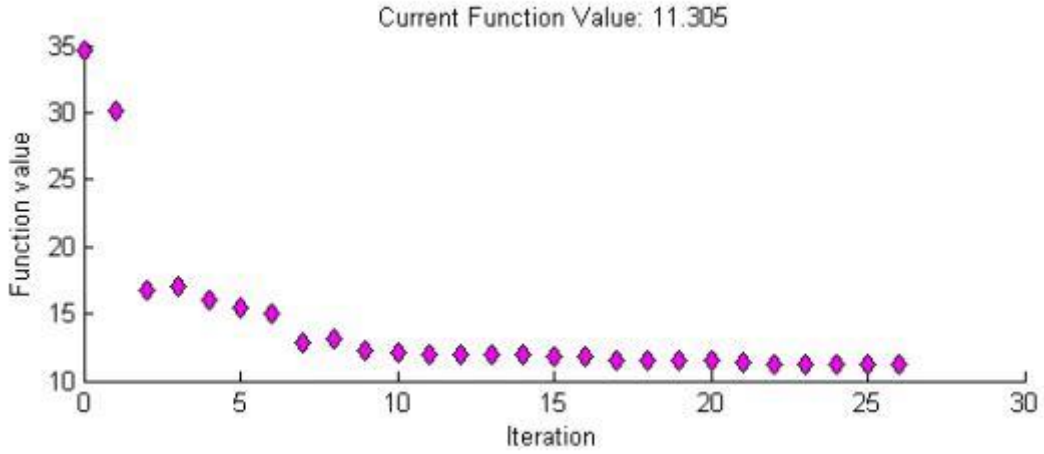


Fig. 3. Actual function value of optimization of zinc plating time for layer thickness of 12 µm

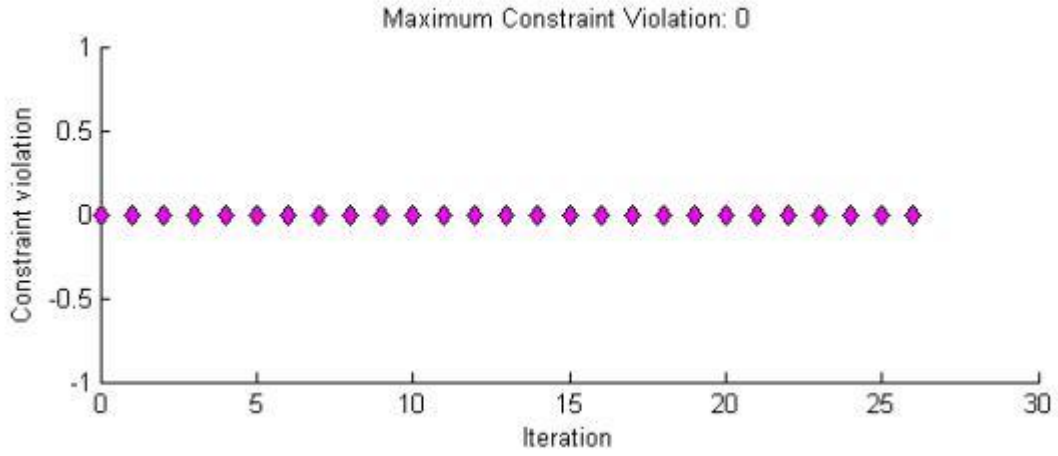


Fig.4. Maximum value of violation of function of optimization of zinc plating time for layer thickness of 12 µm

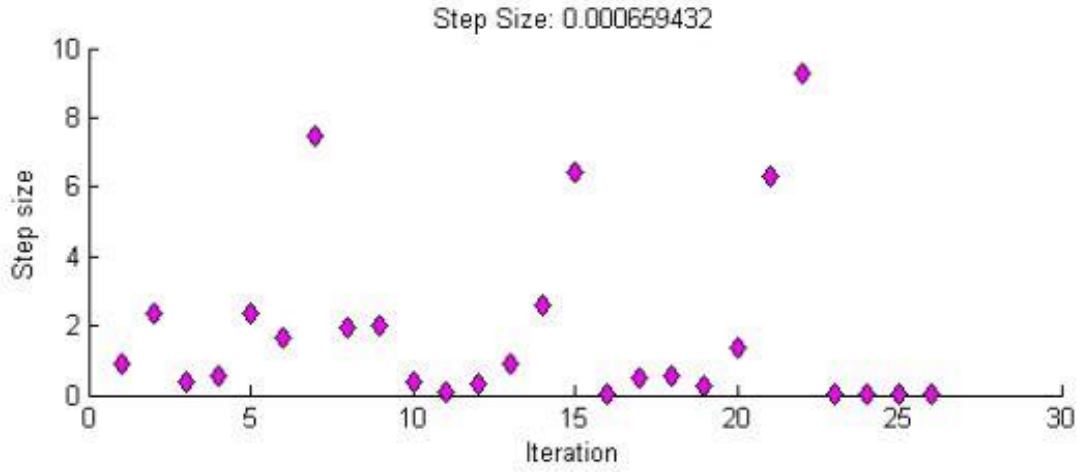


Fig. 5. Step size of optimization of zinc plating time for layer thickness of 12 μm

CONCLUSION

It is not suitable to use COST approach at experimental work at analysis of technological processes of surface treatment, when it is important to observe mutual influence of several functioning factors at the same time [2]. The application of DOE approach is more suitable one, which is presented in the paper that solves a particular problem from practice, where prediction equation set on the basis of statistical analysis of experimentally obtained data in the process of electrolytic alkaline zinc plating was optimized by DOE application and non-linear programming application. The main contribution of the paper is the fact that optimal process conditions of surface treatment were found so that fulfilment of demands of final customers was secured.

REFERENCES

[1] BOX, G. E. P., HUNTER, J. S., HUNTER, W. G. Statistics for Experimenters. 2008, pp. 639. ISBN 978-0-471-71813-0.

[2] ERIKSSON, L. et all. *Design of Experiments – Principles and Applications. Sweden: Umetrics* AB, 2008. ISBN-13: 978-91-973730-4-3.

[3] MELOUN, M., MILITKÝ, J. *Statistická analýza experimentálnich dat*. Praha: ACADEMIA, 2004, 960 pp. ISBN 80-200-1254-0-

[4] YU, J.,WANG, L., JE, L., AI, X., YANG, H. Temperature effects on the electrodeposition of zinc. *Journal of the Electrochemical Society*, vol. 150, 2003, pp. C19–C23. ISSN 00134651.

[5] BALLESTEROS, J.C., D´IAZ-ARISTA, P., MEAS Y., ORTEGA R., TREJO G. Zinc electrodeposition in the presence of polyethylene glycol 20000. *Electrochimica Acta*, vol. 52, no. 11, 2007, pp. 3686–3696. ISSN 0013-4686.

[6] MORÁVKA, J., MAROŠ, B., MICHALEK, K. Vliv neortogonality plánu experimentu na statistickou korektnost modelu. In: *Mezinárodní konference Technical Computing*, Prague 2008, pp. 73. Available at: http://dsp.vscht.cz/konference_matlab/MATLAB08-

[7] HEBÁK, P., HUSTOPECKÝ, J. Vícerozměrné statistické metody s aplikacemi. Praha: SNTL/ALFA, 1987, 456 pp. ISBN 80-01-01076-7.

[8] PHILIP H. DYBVIG. Numerical Methods for Optimization. *Fin500J Mathematical Foundations in Finance*, 2010, 25 pp. [Online].

Acknowledgement
The paper was worked out with the support of APVV SK-CZ-2013-0138 project and VEGA project, No. 1/0738/14.

POSSIBILITIES OF ENSURING PROTECTION OF SELECTED OBJECTS OF CRITICAL INFRASTRUCTURE

Vašková Michaela, Krahulec Josef, Barta Jiří
University of Defence
Kounicova 65, Brno, 60200, Czech Republic
michaela.vaskova@seznam.cz, josef.krahulec@unob.cz,
jiri.barta@unob.cz

Abstract: *With the increasing dependence of countries on the critical infrastructure, it increases their vulnerability. Big threat is primarily in the human factor and especially in terrorist attacks. The biggest breakthrough in the approach to the protection of the critical infrastructure has occurred after September 11, 2001, when there was a terrorist attack in the United States. Based on this event protection of critical infrastructure against terrorist attacks came to the fore. The emphasis is put on the application of the security audit method on the selected objects of the critical infrastructure to find gaps in the critical infrastructure security. The research will be also focused on the crisis preparedness of selected municipalities and results of this research will be used as a foundation for the evaluation of crisis preparedness of critical infrastructure objects in selected areas.*

Keywords: safety audit, critical infrastructure, object of critical infrastructure

INTRODUCTION

Approaches to the protection of the critical infrastructure have been long developing not only at home but also abroad. The biggest breakthrough in has occurred after September 11, 2001, after terrorist attack in the United States. Based on this event, the protection of critical infrastructure came to the fore. To ensure safety of endangered objects of the critical infrastructure by such a terrorist attack, it is appropriate to apply the method of the security audit for identifying the weak points.

1. PROTECTION OF CRITICAL INFRASTRUCTURE

This chapter discusses problematics of critical infrastructure of the Czech Republic and using of the safety audit on selected object of the critical infrastructure.

1.1 Critical infrastructure of the Czech Republic

The basic function of government is to ensure the protection and development of the protected interests and sustainable development of human society. The Constitution of the Czech Republic, as the highest legal document of the Czech Republic, declares that the protected interests of the state are the goals that are cherished as a priority, i.e. the lives and health of people, property, the environment and safety.
Critical infrastructure of the Czech Republic is defined as production and non-production systems and services, whose malfunction could have a serious impact on national security, the economy, public administration and on ensuring of fundamental life needs

of the population. [1]

The object of critical infrastructure is then defined as a building or facility to ensure the functioning of critical infrastructure. Objects of critical infrastructure are the production and non-production systems and services whose disruption or complete destruction would have a serious impact on the running of the state, for its operations and performance of its functions. [1]

1.2 Selected objects of the critical infrastructure

As one of the objects of critical infrastructure we chose, as a model, airport Brno-Tuřany. There will be carried out a research focused on the airport of Vaclav Havel in Prague, in the near future. As a second object of the critical infrastructure, there was chosen the Nuclear Power Plant Dukovany.

To enhance the protection of the objects of the critical infrastructure and minimize the risk of attacking those objects, it is appropriate to apply the security audit method to find weak points (gaps) in security.

1.3. Using of security audit method

Security audit is a systematic, if it is possible, independent examination to determine whether all activities and related to results comply with planned arrangements and whether these arrangements are implemented effectively and if they are suitable to achieve objectives and policies of the organization. Audit is an integral part of security management. It is a very effective tool to check its status and the status of the entire organization. [2]

Audit is an independent, documented process that aims to determine whether activities and related results comply with audit criteria, and to what extent. Audit criteria may be procedures and requirements of the organization, procedures, politics etc. The outcome of the security audit is not only the assessment of compliance, but also to assess the effectiveness and reliability of safety management. The audit must take into account:

- Effectiveness of the organization,
- Risks,
- Level control and process efficiency,
- Level of management and process efficiency,
- Opportunities for cost reduction, waste and other forms of waste,
- Opportunities for process improvement, the overall security status of the organization. [2]

To make audit plan to function, it is necessary to pay great attention to the selection and qualification of auditors. Procedures for carrying out the audits a company prepares itself and in accordance with the standard must include:

- The subject and scope of the audits and their frequency,
- Auditing methodology, defining responsibility and authority for the audit program, the audit arrangements in terms of management,
- Own auditing procedure,

- The conditions and specifications to present reports on the results of the audit,
- Competence requirements and training of auditors,
- Own auditing procedure,
- The conditions and specifications to present reports on the results of the audit,
- Competence requirements and training of auditors,
- Way to discuss the audit findings with relevant staff,
- The monitoring and verification of the effectiveness of corrective measures. [2, 3, 4]

With the connection of ensuring the protection with the help of security audit, there is also necessary to focus the attention on critical infrastructure protection but also on crisis preparedness of municipalities, because there are many activities of security protection closest to the citizens.

2. EMERGENCY PREPAREDNESS OF MUNICIPALITIES

According the current legislative framework there is established new legislative background concerning with municipalities with extended power. Within the scope of ensuring security as it was mentioned before moves the attention to occupy with protection on municipality level, concretely said crisis preparedness to deal with extraordinary events. The branch of population protection and crisis management is sophisticated and it is really important to put the attention not only on crisis response system from the point of view of law but also from the point of view of municipality bodies. It means to concentrate on the role of municipality managers they are closest to their citizens and deal with extraordinary event just on the hot spot, where this situation happened. According the interview with professionals from crisis management departments of municipalities, they are saying that current situation about crisis preparedness is everything prepared but some problems could appear when new elected chiefs of municipalities are not so educated and erudite to be able to deal with crisis or another situation. From the point of view of ensuring the municipality protection is really important to establish a team of people that will be appropriate educated and be able to make the preparedness background of municipality to deal with mentioned events. [5]

Also necessary was to establish some criteria that would be able to evaluate crisis preparedness of critical infrastructure elements because this problem has not yet been modified in the czech law. Indistinct competences and relationships were the basic of establishing new legislative framework to define specific rights and obligations of crisis management of the municipality. Concrete and specific informations is possible to find in the law of crisis management 240/2000 Coll., § 10, para. 1., about coordination the preparation for crisis situations and their solutions the Ministry of Interior. The part of this Ministry is General Directorate of Fire and Rescue Service of the Czech Republic. [5, 6]

This is the analysis part of the research and in future research there will be the attention focused on crisis preparedness of municipalities, especially municipalities with extended powers. The major aim is to declare and evaluate the situation how municipalities in the Czech Republic are prepared to deal with crisis situations and extraordinary events. [7, 8] The fulfil approach will contain particular parts. There will be the current status analysis of problem concerning to the main topic, how the municipalities are prepared in these days, what they need and where we can find potential problems. After this research part will be performed the analysis part of the research with the main aim to define current readiness of municipalities with extended powers, especially with emphasis on present status where data

and information are collected by using interview and surveys, also from received materials concerning crisis management. [6, 8]
As a technical and software support of the research will be used sophisticated software and programs with integrated database which will be able to provide and model an adequate schema of actors, entities and environments during dealing with extraordinary events. There will be used KISKAN program, which uses database system for supporting crisis management processes and also ensures business continuity. It is able to provide processing of crisis plans, emergency plans and plans of crisis preparedness, also ensures information exchange about readiness to deal with crisis situations and extraordinary events among independent subjects. KISKAN program focuses on the environment in which information can be processed in accordance with provisions of the Act No. 240/2000 Coll., on crisis management and Act No. 239/2000 Coll., on integrated rescue system. KISKAN is able to support processes such as: risk assessment, local and distant crisis situations solution, sources preparation, measures planning, document creating many others. Among the main functions for supporting crisis management processes of KISKAN belong: processing of overview possible risks, sources integration of all information for crisis readiness into one relational database, creating of connecting overview to the crisis management subjects, accounting, to specify planned measures on the basis of experience with crisis situations, creating a centre for receiving SMS messages and sending notification, GPS surveillance positions and routes of mobile resources in real time, local and remote activation of a crisis situation by an activation code and monitoring their performance in real time, secure data exchange electronic signature and encryption, automated updates of the plan dealing with the crisis according to the status of tasks, further processing of the data in Microsoft Word and Microsoft Excel, synchronization information from remote databases and many others. [6, 9]

All collected information will be processed together into specific database which will evaluate them and by using this programme is possible to compare answers of responsible persons, their relations etc. display on one scene. As a result will be concrete recommendations how to process and evaluate current status of crisis preparedness, how to find some shortages and suggest new approaches to improve the area of crisis management. [7, 9]

3. SIMULATION PROGRAMS FOR TESTING EMERGENCY PLANS

Simulation is an imitation of some real thing, condition or process. The act of simulation of something itself generally means displaying some key features or behaviour of selected physical or abstract systems. Simulation is used in many contexts comprising modelling of natural or human systems with the aim to obtain knowledge about their behaviour. [10] Other contexts comprise technological simulations for optimizing the performance, security engineering, testing, training and educating. Simulation can be used for visualisation of possible real impacts, alternative conditions and ways of acting. Key issues in simulation comprise e.g. obtaining valid sources of information about corresponding selection of key characteristics and behaviour, using the simplifying estimation and prerequisites in the frame of simulation as well as reliability and validity of the results of the simulation given. [6, 11]

For the training preparation and verifying crisis plans, instructors can use various computer programs which enable better graphic visualisation of the solution, practice different ways of dealing with the different situations and the way of command. What is more, they can represent a tool for the various roles in the process of solution of the emergency situation. The environment of these programs increases the effect of preparation, which results in being more realistic and the trainees will better memorise the trained actions. [8] To verify crisis

plans and crisis staffs of personnel currently we use the program One Semi-Automated Forces, which was developed for the army, but after some modification it is also possible to use it in civil sector.

3.1 One Semi-Automated Forces

Simulator OneSAF is a program of constructive simulation already used for several years which became widely used according to the needs of training as well as requirements of the trainees. Currently, a wide spectrum of CAX training types can be carried out by it. The program has been further adapted and adjusted according to the needs of the Army of the Czech Republic especially in relation with introducing new armament and equipment. [10]

System has been extended by the elements of the Integrated Rescue System IRS and is widely used for the training of crisis management staff/specialists of IRS, especially HZS and PCR.
It is an older simulation system of constructive simulation. Nowadays its technology is outdated. It is primarily aimed at military purposes. Simulation of the activities of the units of IRS is feasible only partially with certain restrictions. From this reason is currently implemented simulate system WASP-C, which simulate extraordinary events and activities of forces and means of the integrated rescue system (IRS) and other players in real time.

3.2 Simulation system WASP

It represents a system of constructive simulation for the computerised generation of forces and creation of synthetic environment. Originally it was designed for the use of army but the version for the components of the Integrated Rescue System called WASP-C has been developed as well. The simulator enables to practice management on the tactical, operational and strategic levels. Modelling of various emergency situations and their solution is possible in this environment.

Environment in the simulator is ensured by the combination of terrain database created from the detailed geographical data, model of weather and other dynamic environmental models. Terrain database contains all common objects in the countryside (bodies of water, roads, built-up areas, vegetation, relief, type of soil and other objects). Individual objects have predefined features influencing simulation of their own entities in relation to their purpose. Weather editor enables to set basic parameters (date and time, air temperature, velocity and direction of the wind, type and intensity of precipitations, humidity and pressure of air, type of cloud cover, light intensity etc.). Some of the parameters are mutually interlinked based on the actions happening in the atmosphere known from meteorology. Dynamic models of environment enable to modify the countryside with objects and phenomena which can change their form in the course of time. There are accidents simulated in great detail as well as a vast database of forces and means. Program puts more emphasis on the correct execution than on graphic output and it is aimed at the group of trainees as well as at an individual. [12]

Concept of the program is suitable for the use in practical training of solving emergency events with the mutual cooperation of the intervening units. The system is completed by a communication system Astra, which simulates normal means of communication (telephones, radios, PTT, etc.). Exercise and verification of crisis plans in an environment and only with funds that have crisis teams routinely available.

This simulator has proven in previous practical exercises, especially when exercising in Hustopeče, where the exercise was carried out to verify the crisis plans of the municipality. Exercise was attended by surrounding municipalities, the IRS and other stakeholders. [12] Conclusions from the exercise helped unify procedures in dealing with similar incidents and to give impetus to further cooperation between the municipality Hustopeče, local companies and the IRS.

CONCLUSION

The current situation raises a claim for continual improvement of safety relative to existing as well as future threats. Emphasis is put mainly on education and erudition managers, but also to implement the latest technologies and practices that can contribute to improving safety, not only in terms of protection of critical infrastructure, but also in terms of ensuring the protection of the population and improving the crisis management process.

The solution of crisis situation can be designed and provided by the crisis continuity scenarios and sophisticated methodological approach. The newest requirements to provide protection of municipalities, its population and property is still significant item and especially for municipalities, because after legislative reform is this branch full of gaps that can be studied.

The actual simulation cannot replace the practical deployment in emergencies where trespassing gain unparalleled practical experience and crisis plans are proved in practice. However, for the purpose of preventing and preparing for emergencies and crisis situations is the best practical training. [11] In artificially induced emergencies through constructive simulator, workers can check not only emergency staff contingency plans, but also their own communication and leadership skills.

REFERENCES

[1] Kritická infrastruktura – Ministerstvo vnitra České republiky. MINISTERSTVO VNITRA ČESKÉ REPUBLIKY. *Úvodní strana – Ministerstvo vnitra České republiky* [online]. ©2015 [cit. 2015-09-21]. Dostupné z: http://www.mvcr.cz/clanek/kriticka-infrastruktura.aspx

[2] ŠEBESTOVÁ, Marie; STANĚK, Miroslav. *Komentované vydání ČSN EN ISO 19011:2003 Směrnice pro auditování systému managementu jakosti a/nebo systému environmentálního managementu*. Praha: Český normalizační institut, 2003, 74 s. ISBN 80-7283-112-7.

[3] RINGEL, Miroslav. *Bezpečnostní audit v průmyslovém podniku. Brno,* 2009. Diplomová práce. Vysoké učení technické v Brně. Vedoucí práce Babinec.

[4] *What are safety audit? SASA* [online]. *The society of accredited safety auditors limited,* ©1999 [cit. 2015-11-02]. Dostupné z: *http://www.sasa.org.hk/audit.htm#5.*

[5] KOČÍ, R. *Obecní samospráva v České republice: praktická příručka s judikaturou.* Vol. 1. Praha: Leges, 2012, 240 p. Praktik. ISBN 978-808-7576-281.

[6] URBÁNEK, J. F. et al. Crisis Scenarios. Brno: Univerzity of Defence, 2013. 240 pp. ISBN: 978-80-7231-934-3.

[7] WATTERS, J. *Disaster Recovery, Crisis Response and Business Community A Management Desk Reference.* Apress, 2013. ISBN 978-143-0264-064.

[8] ŘEHÁK D, GRASSEOVÁ M. The ways of assessing the security of organization information systems through SWOT analysis, pp. 162-184. DOI: 10.4018/978-1-61350-311-9.ch007. In ALSHAWI, Mustafa, ARIF, Mohammed (eds.). Cases on E-Readiness and Information Systems Management in Organizations: Tools for Maximizing Strategic Alignment. 1st edition. Hershey, PA, USA: IGI Global, 2011. 318 p. ISBN 978-1-61350-311-9. DOI: 10.4018/978-1-61350-311-9.

[9] SVOBODA, Ivo a Karel SCHELLE. *Základy organizace veřejné správy*. Vyd. 1. Ostrava: Key Publishing, 2006, 206 s. Právo (Key Publishing). ISBN 80-239-8011-4.

[10] KOVÁŘÍK, František. *112 odborný časopis požární ochrany, integrovaného záchranného systému a ochrany obyvatelstva: Cvičení krizového štábu v Poličce.* Praha: MV - generální ředitelství HZS ČR, 2015, 2015(8). ISSN 1213-7057.

[11] REHAK D, SENOVSKY P. Preference Risk Assessment of Electric Power Critical Infra-structure. Chemical Engineering Transactions, 2014, Vol. 36, pp. 469-474. ISSN 1974-9791. DOI: 10.3303/CET1436079.

[12] *VR Group: Mobile Crisis Management Training System* [online]. Brno, 2014 [cit. 2015-10-14]. Dostupné z: http://www.vrgroup.cz/index.php/products-solutions/crisis-management/emergency-committee-training

Acknowledgement

The work presented in this paper has been supported by the project by Technology Agency of the Czech Republic with the topic Research and Development of Simulation Instruments for Interoperability Training of Crisis Management Participants and Subjects of Critical Infrastructure (research project No. TA04021582) and results presented in this article are also obtained in the project called The use of simulation methods and modelling system to ensure the continuity of the organization in terms of societal security (project code SV14-FEM-K106-07-KRA).

REMARKS ON COMPACT SUBMEASURES

Tomáš Visnyai
Slovak University of Technology in Bratislava
Faculty of Chemical and Food Technology
Radlinského 9, 812 37 Bratislava, Slovakia
tomas.visnyai@stuba.sk

Abstract: *In this paper, we formulate a generalization of a sufficient condition for the convergence of series with positive terms published by Estrada and Kanwal (1986). We give a necessary and sufficient condition to such a density for a certain class of matrices be a compact submeasure. Further we provide an example of the regular matrix for which the density defined by this matrix is not compact submeasure. Finally, an exponential density of sets is defined and it is shown that it is not a compact submeasure whenever if the set* $A \subseteq \mathbb{N}$.

Keywords: convergence of series, density of sets, compact submeasure

1. INTRODUCTION

In the papers [9,10,12] the notion of compact submeasure was introduced. The set function $m: 2^{\mathbb{N}} \to \langle 0, +\infty)$ is called a submeasure if it is monotone and subadditive, i.e.

i) $A \subseteq B \Rightarrow m(A) \leq m(B)$
ii) $m(A \cup B) \leq m(A) + m(B)$

The submeasure m is called compact if

iii) $m(\{a\}) = 0$ for every $a \in A$
iv) for every $\varepsilon > 0$ there exists a decomposition $\mathbb{N} = A_1 \cup A_2 \cup \ldots \cup A_s$ of $\mathbb{N}$ such that $m(A_j) < \varepsilon$ for each $j = 1,2, \ldots s$.

Before we define the concept of density recall the concept of a regular matrix. A method defined by the infinite matrix $\mathbb{T} = (a_{nk})$ $n, k = 1,2 \ldots$ is said to be regular if for all convergent sequences $x = (x_k)$ for which $\lim_{k\to\infty} x_k = L \in \mathbb{R}$ implies that the sequence $t_n = \sum_{k=1}^{\infty} a_{nk} x_k$ converges to $L \in \mathbb{R}$. It is well-known that the matrix $\mathbb{T} = (a_{nk})$ is regular if and only if it satisfies the following three conditions (see [11]):

a) $\exists\, M > 0, \forall n = 1,2, \ldots \ \sum_{k=1}^{\infty} |a_{nk}| \leq M$
b) $\lim_{n\to\infty} a_{nk} = 0 \ k = 1,2, \ldots$
c) $\lim_{n\to\infty} \sum_{k=1}^{\infty} a_{nk} = 1.$

For example, Caesaro-matrix $C = (c_{nk})$, where $c_{nk} = \frac{1}{n}, k \leq n$; $c_{nk} = 0, k > n$ is regular.

Definition 1.1.

Let $\mathbb{T} = (a_{nk})$ be a nonnegative regular matrix and $A \subseteq \mathbb{N}$. Let

$$d_{\mathbb{T}}^{(n)}(A) = \sum_{k=1}^{\infty} a_{nk}\chi_A(k)\,, n = 1,2, \ldots,$$

where χ_A being characteristic function of A. Then

$$\bar{d}_{\mathbb{T}}(A) = \lim_{n\to\infty} \sup d_{\mathbb{T}}^{(n)}(A) \quad \text{is called upper } \mathbb{T}\text{- density of } A$$

and

$$\underline{d}_{\mathbb{T}}(A) = \lim_{n\to\infty} \inf d_{\mathbb{T}}^{(n)}(A) \text{ is called lower } \mathbb{T}\text{- density of } A.$$

If $\underline{d}_{\mathbb{T}}(A) = \bar{d}_{\mathbb{T}}(A) = d_{\mathbb{T}}(A)$, then $d_{\mathbb{T}}(A) = \lim_{n\to\infty} \sum_{k=1}^{\infty} a_{nk}\chi_A(k)$ is a $\mathbb{T}$- density of A.
By the regularity of $\mathbb{T} = (a_{nk})$ it is clear that $d_{\mathbb{T}}(A) \in \langle 0,1\rangle$.

Below are some examples.

Example 1.2.

Let $\mathbb{T}_Z = (z_{nk})$, where $z_{n,k} = z_{n,k+1} = \frac{1}{2}$ if $k = n$ and $z_{n,k} = 0$ otherwise. Matrix $\mathbb{T}_Z = (z_{nk})$ is regular and called Zweier matrix. It is easy to see, that $d_{\mathbb{T}_Z}(A) \in \{0, \frac{1}{2}, 1\}$ if it exists.

Example 1.3.

Let $\mathbb{T}_W = (a_{nk})$ is regular matrix defined by following way:

$$a_{nk} = \frac{c_k}{S(n)} \qquad \text{for } k \le n,$$

$$a_{nk} = 0 \qquad \text{for } k > n,$$

where $c_n > 0\ (n = 1,2, \ldots), \sum_{n=1}^{\infty} c_n = +\infty, S(n) = c_1 + \ldots + c_n$. Then $d_{\mathbb{T}_W}(A) = \lim_{n\to\infty} \frac{1}{S(n)} \sum_{k=1}^{\infty} c_k\chi_A(k)$. Specifically, if $c_n = 1$ for every $n \in \mathbb{N}$, then we get $d_{\mathbb{T}}(A) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{\infty} \chi_A(k) = d(A)$ asymptotic density of the set A. If $c_n = \frac{1}{n}$ for every $n \in \mathbb{N}$, then we get $d_{\mathbb{T}}(A) = \lim_{n\to\infty} \frac{1}{\ln n} \sum_{k=1}^{\infty} \frac{1}{k}\chi_A(k) = \delta(A)$ logarithmic density of the set A (see [6,7,8,10]). If $c_n = n^{\alpha}$ for every $n \in \mathbb{N}$ and $\alpha \in (0,1)$ the matrix $\mathbb{T}_W = (a_{nk})$ is Riesz matrix of type k^{α}(cf.[4]). Generally, density defined by the matrix $\mathbb{T}_W$ is called weighted density of the set $A \subseteq \mathbb{N}$ (cf.[10]).

Example 1.4.

Let $\mathbb{T}_N = (a_{nk})$ is a regular matrix defined by the following way:

$$a_{nk} = \frac{c_{n-k+1}}{S(n)} \qquad \text{for } k \leq n,$$

$$a_{nk} = 0 \qquad \text{for } k > n,$$

where $c_n > 0\ (n = 1,2,\dots), \sum_{n=1}^{\infty} c_n = +\infty, S(n) = c_1 + \dots + c_n$.
Then $d_{\mathbb{T}}(A) = \lim_{n\to\infty} \frac{1}{S(n)} \sum_{k=1}^{\infty} c_{n-k+1} \chi_A(k)$ is a density defined by at the Norlund matrix. It is known that Norlund matrix is regular if and only if $\lim_{k\to\infty} \frac{c_k}{S(k)} = 0$ (see [11]).

Example 1.5.

According to the Steinhaus theorem ([11, Lemma 3.5.4.]) for every regular matrix there exists a sequence of 0's and 1's which is not summable by this matrix. Such a sequence is the characteristic function of a any set. Hence there is a set $B \subseteq \mathbb{N}$ which has not a $\mathbb{T}$- density.

Finally, we give two type of densitites which can not be defined by a regular matrix.
Let $A \subseteq \mathbb{N}$. The upper uniform density term $\bar{u}(A) = \lim_{n\to\infty} \left[\max_{m\geq 0} \frac{1}{n} \sum_{i=m+1}^{m+n} \chi_A(i)\right]$ and lower uniform density term $\underline{u}(A) = \lim_{n\to\infty} \left[\min_{m\geq 0} \frac{1}{n} \sum_{i=m+1}^{m+n} \chi_A(i)\right]$. Where it is equal to their common value is $u(A)$ uniform density of A. (see [1,3]). The upper and lower exponential densities of an infinite subset $A \subseteq \mathbb{N}$ are defined by $\bar{\varepsilon}(A) = \lim_{n\to\infty} \sup \frac{\ln \sum_{k=1}^{n} \chi_A(k)}{\ln n}$, $\underline{\varepsilon}(A) = \lim_{n\to\infty} \inf \frac{\ln \sum_{k=1}^{n} \chi_A(k)}{\ln n}$ respectivelly. If $\bar{\varepsilon}(A) = \underline{\varepsilon}(A)$ then we say that A has the exponential density. In case when $A = \{a_1 < a_2 < \cdots\}$ is a infinite then $\bar{\varepsilon}(A) = \tau(A) = \lim_{n\to\infty} \frac{\ln n}{\ln a_n}$ is an exponent of convergence of the sequence (a_n) (see [5]).

2. MAIN RESULTS

In 1986 Estrada and Kanwal proved that if a series with positive terms converges along each set of the zero asymptotic density then it converges in the usually sense as well. It means that a series with positive terms is divergent there is a set $B \subseteq \mathbb{N}$ with zero asymptotic density also that the series divergent along this set (cf.[2]). For example the harmonic series is divergent hence there is a set $\mathbb{P}$ of all primes having zero asymptotic density and series of reciprocal values primes is divergent too. M. Paštéka generalized this result. He replace the term asymptotic density with compact submeasure.(see [9]).

We denote $\mathcal{I}_{\varrho} = \{A \subseteq \mathbb{N}: \bar{\varrho}(A) = 0\}$ class the subsets of $\mathbb{N}$, where $\bar{\varrho}$ is arbitrary density defined in this article. Following inclusion is true:

$$\mathcal{I}_u \subsetneq \mathcal{I}_d \subsetneq \mathcal{I}_{\delta} \subsetneq \mathcal{I}_{\mathbb{T}}.$$

For example set $B = \bigcup_{n=1}^{\infty}\{n^3+1, \dots n^3+n\}$ have not uniform density, but its asymptotic density is equal to zero. Further the set $C = \bigcup_{n=1}^{\infty} C_n$, where $C_n = \left\{n^{n^2}+1, \dots, n^{n^2+1}\right\}$ have not asymptotic density but its belong to the class $\mathcal{J}_\delta$.

In [12] it is proved this Theorem:

Theorem 2.1.

Let $\sum_{k=1}^{\infty} a_k$ be a series with positive terms. If for each $A \subseteq \mathbb{N}$ with $\bar{u}(A) = 0$ we have $\sum_{k\in A} a_k < +\infty$, then $\sum_{k=1}^{\infty} a_k < +\infty$.

Proof.
It can be easily cheeked that $\bar{u}$ satisfies properties i) – iv) of the compact submeasures. Since $\varepsilon > 0$ then we choose an $m \in \mathbb{N}$ such that $\frac{1}{m} < \varepsilon$. The desired decomposition of $\mathbb{N}$ can be taken by decomposition of the residual class, e.g. $\mathbb{N} = \bar{0} \cup \bar{1} \cup \dots \cup \overline{(m-1)}$. It is true that $\bar{u}\overline{(m-1)} = \frac{1}{m}$ (see [9]). Hence $\bar{u}$ is a compact submeasure on $2^{\mathbb{N}}$.

■

We ask yourself a natural question: when the $\mathbb{T}$- density is a compact submeasure? For the upper density $\bar{d}_{\mathbb{T}}$ defined by the matrix $\mathbb{T}_W = (a_{nk})$ (see Example 1.3.) we find a necessary and sufficient condition such that to be a compact submeasure on $2^{\mathbb{N}}$.

Theorem 2.2.

Let $\mathbb{T}_W = (a_{nk})$ is a regular matrix defined in Example 1.3.
Then $\bar{d}_{\mathbb{T}_W}(A) = \lim_{n\to\infty} \sup \frac{1}{S(n)} \sum_{k=1}^{\infty} c_k \chi_A(k)$ is a compact submeasure if and only if $\lim_{n\to\infty} \frac{c_n}{S(n)} = 0$.

Proof.
Let $\lim_{n\to\infty} \frac{c_n}{S(n)} = 0$ is holds. According to Theorem 1.2. in [8] the upper density $\bar{d}_{\mathbb{T}_W}$ has Darboux property. It follows that $\mathbb{N}$ can be decomposed into $\mathbb{N} = \mathbb{N}_1^{(1)} \cup \mathbb{N}_2^{(2)}$ such that $\bar{d}_{\mathbb{T}_W}\left(\mathbb{N}_1^{(1)}\right) = \frac{1}{2} = \bar{d}_{\mathbb{T}_W}\left(\mathbb{N}_2^{(2)}\right)$. In this way we can construct by induction a decomposition $\mathbb{N} = \mathbb{N}_1^{(k)} \cup \dots \cup \mathbb{N}_{2^k}^{(k)}$ such that $\bar{d}_{\mathbb{T}_W}\left(\mathbb{N}_j^{(k)}\right) = \frac{1}{2^k}$, $j = 1,2,\dots,2^k, k = 1,2,\dots$, and this implies that $\bar{d}_{\mathbb{T}_W}$ is a compact submeasure.

Let $\lim_{n\to\infty} \frac{c_n}{S(n)} = 0$ do not hold. Then there exists an infinite sequence (n_k) and $\alpha > 0$ such that $\frac{c_{n_k}}{S(n_k)} > \alpha$, $k = 1,2,\dots$. Consider a decomposition $\mathbb{N} = A_1 \cup A_2 \cup \dots \cup A_s$. This

decomposition is finite, therefore one of the sets $A_1, \dots, A_s$ must contain infinitely many elements of (n_k). Let $A_m = \{n_{k(1)}, n_{k(2)}, \dots\}$. Then $d_{\mathbb{T}_W}^{(n_{k(j)})} \geq \frac{c_{n_{k(j)}}}{S(n_{k(j)})} > \alpha$ and so $\bar{d}_{\mathbb{T}_W}(A_m) \geq \alpha$. Therefore $\bar{d}_{\mathbb{T}_W}$ is not a compact submeasure. ■

Corolarry 2.3.

a) From Theorem 2.2. follows, that upper asymptotic and upper logarithmic density are compact submeasure on $2^{\mathbb{N}}$.

b) Thus in Theorem 2.1. can by replaced the upper uniform density $\bar{u}$ by the density $\bar{d}_{\mathbb{T}_W}$ if $\lim_{n\to\infty} \frac{c_n}{S(n)} = 0$ holds.

Theorem 2.4.

There exists a regular matrix $\mathbb{T}_0 = (a_{nk})$ for which the upper density $\bar{d}_{\mathbb{T}_0}$ is not compact submeasure.

Proof.

Let us put $c_n = n^n$,$n = 1,2, \dots$ in Example 1.3. Subsequently $a_{nk} = \frac{k^k}{1+2^2+\dots+n^n}$, $k \leq n$ and $a_{nk} = 0$ otherwise. It is easy to see that $\mathbb{T}_0 = (a_{nk})$ is regular but condition of Theorem 2.2. is not satisfied: $\lim_{n\to\infty} \frac{c_n}{S(n)} = \lim_{n\to\infty} \frac{n^n}{1+2^2+\dots+n^n} = \lim_{n\to\infty} \frac{\frac{1}{n}}{\int_0^1 x^n dx} = 1.$

■

Finally, we prove that the upper exponential density is not compact submeasure. In [5] it is shown that for the set $A = \{a_1 < a_2 < \cdots\}$ we have $\bar{\varepsilon}(A) = \tau(A) = \lim_{n\to\infty} \frac{\ln n}{\ln a_n}$ where

$$\tau(A) = \inf\left\{t > 0 : \sum_{k=1}^{\infty} a_k^{-t} < +\infty\right\}$$

is an exponent of convergence. It is unknown that for $A, B \subseteq \mathbb{N}$ it holds $\tau(A \cup B) = \max\{\tau(A), \tau(B)\}$ (see [12]).

Theorem 2.5.

Let $A = \{a_1 < a_2 < \cdots\} \subseteq \mathbb{N}$. Then upper exponential density $\bar{\varepsilon}(A)$ is not compact submeasure.

Proof.

We assume that $\bar{\varepsilon}(A)$ is a compact submeasure, $A = \{a_1 < a_2 < \cdots\}$. On the basis of iv) properties of compact submeasure for every $\epsilon > 0$ there exists a decomposition $\mathbb{N} = A_1 \cup A_2 \cup \ldots \cup A_s$, also that $\bar{\varepsilon}(A) < \epsilon, j = 1,2,\ldots,s$. Let $0 < \epsilon < 1$.

Then $1 = \bar{\varepsilon}(\mathbb{N}) = \tau(\mathbb{N}) = \tau(A_1 \cup A_2 \cup \ldots \cup A_s) = \max_{1 \le j \le s}\{\tau(A_1), \ldots, \tau(A_s)\} \le \epsilon < 1.$

This is contradiction. ■

Consequently can not replace $\bar{u}(A)$ with $\bar{\varepsilon}(A)$ in the Theorem 2.2.

Open problem

Is a density defined by Norlund matrix (Example 1.4.) compact submeasure?

REFERENCES

[1] BALÁŽ, V., ŠALÁT, T. *Uniform density u and corresponding $\mathcal{I}_u$ - convergence*, Math. Commun., 11, (2006), pp. 1-7.

[2] ESTRADA, R., KANWAL, R.P. *Series that converge on sets of null density.* Proc.Amer.Math.Soc., 97, (1986), pp. 682-686.

[3] FREEMAN, A.R., SEMBER, J.J. *Densities and summability*. Pacific Journal of Math., Vol.95, No.2, (1981), pp. 293-305.

[4] GOGOLA, J., MAČAJ, M., VISNYAI, T. *On $\mathcal{I}_c^{(q)}$- convergence*. Annales Mathematicae et Informaticae, 38, (2011), pp. 27-36.

[5] GREKOS, G., SLEZIAK, M. TOMANOVÁ, J. *Gaps and the exponent of convergence of an integer sequence*. Uniform Distribution Theory 6, No.2, (2011), pp. 95-116.

[6] KOSTYRKO, P., MAČAJ, M., ŠALÁT, T., SLEZIAK, M. *$\mathcal{J}$- convergence and extremal $\mathcal{J}$- limit points*. Math. Slov. 55, No.4, (2005), pp. 443-464.

[7] KOSTYRKO, P., ŠALÁT, T. WYLCZYŃSKI, W. *$\mathcal{J}$- convergence*. Real Analysis Exchange 26, (2000-2001), pp. 669-686.

[8] MAČAJ, M., MIŠÍK, L., ŠALÁT, T., TOMANOVÁ, J. *On a class of densities of sets of positive integers, Acta Math*. Univ. Comenianae, LXXII(2), (2003), pp. 213 - 221.

[9] PAŠTÉKA, M. *Convergence of series and submeasures of the set of positive integers*. Math. Slov. 40, (1990), pp. 273-278.

[10] PAŠTÉKA, M., ŠALÁT, T., VISNYAI, T. *Remarks on Buck`s measure density and a generalization of asymptotic density.* Tatr. Mt. Math. Publ., 31, (2005), pp. 87-101.

[11] PETERSEN, G. M. *Regular matrix transformations*. London, (1966)

[12] ŠALÁT, T., VISNYAI, T. *Subadditive measures on $\mathbb{N}$ and the convergence of series with positive terms*. Acta Math., 6, (2003), pp. 43-52.