Mathematics, Information Technologies and Applied Sciences 2021

post-conference proceedings of extended versions of selected papers

Editors:

Jaromír Baštinec and Miroslav Hrubý

Brno, Czech Republic, 2021

Aims and target group of the conference:

The conference **MITAV 2021** is the eighth annual MITAV conference. It 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 groups, 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 of teachers from different kinds of schools and orientation. We also welcome presentations by (diploma and doctoral) students and teachers who are just beginning their careers, as their novel 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, 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 17 and 18, 2021.

Due to measures against the spread of coronavirus and covid-19, the prezence form of the conference was not possible. The conference was organized as the online conference, the software MS Teams was used.

Conference languages:

English, Czech, Slovak

International Scientific committee:

Prof. Leonid BEREZANSKY
Ben Gurion University of the Negev, Beer Sheva, Israel

Prof. Zuzana DOŠLÁ
Masaryk University in Brno, Faculty of Science, Czech Republic

Prof. Irada DZHALLADOVA
Kyiv National Economic Vadym Getman University, Ukraine

Prof. Mihály PITUK
University of Pannonia, Faculty of Information Technology, Veszprém, Hungary

Prof. Ľubica STUCHLÍKOVÁ
Slovak University of Technology in Bratislava, Faculty of Electrical Engineering and Information Technology, Slovakia

International Programme committee:

Chair: **Jaromír BAŠTINEC**
Brno University of Technology, Faculty of Electrical Engineering and Communication, Brno, Czech Republic

Members:
Jan HODICKÝ
M&S Technical SME at NATO HQ SACT, Norfolk, Virginia, USA

Miroslav HRUBÝ
University of Defence, Faculty of Military Technology, Brno, Czech Republic

Edita KOLÁŘOVÁ
Brno University of Technology, Faculty of Electrical Engineering and Communication, Brno, Czech Republic

Piet KOMMERS
Professor of UNESCO Learning Technologies, Emeritus University of Twente, Enschede, the Netherlands

Nataliia MORZE
Borys Grinchenko Kiev University, Kiev, Ukraine

Václav PŘENOSIL
Masaryk University, Faculty of Informatics, Brno, Czech Republic

Magdalena ROSZAK
Poznan University of Medical Sciences, Department of Computer Science and Statistics, Poznań, Poland

Eugenia SMYRNOVA-TRYBULSKA
University of Silesia, Katowice – Cieszyn, Poland

Olga YAKOVLEVA
Herzen State Pedagogical University of Russia, St. Petersburg, Russia

National steering committee:

Chair: **Karel Lepka**
Masaryk University in Brno, Faculty of Education, Department of Mathematics

Members:
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

Milan Jirsa
University of Defence in Brno, Faculty of Military Technology, Department of Informatics and Cyber Operations

Tomáš Ráčil
University of Defence, Faculty of Military Technology, Department of Informatics and Cyber Operations

Each MITAV 2021 participant received printed collection of abstracts **MITAV 2021** with ISBN 978-80-7582-380-9. CD supplement of this printed volume contains all the accepted contributions of the conference.

Now, in autumn 2021, this **post-conference proceedings** were published, containing extended versions of selected MITAV 2021 contributions. The proceedings are published in English and contain extended versions of 11 selected conference papers. Published articles have been chosen from 17 conference papers and every article was once more reviewed.

Webpage of the MITAV conference:

https://mitav.unob.cz

Content:

List of reviewers:

Leonid Berezansky, Ben Gurion University of the Negev, Beer Sheva, Israel
Irada Dzhalladova, Kyiv National Economic Vadym Getman University, Ukraine
Michal Fusek, Brno University of Technology, Czech Republik
Jiří Jánský, University of Defence, Brno, Czech Republik
Denys Khusainov, Kiev National University, Ukraine
Martin Kovár, Brno University of Technology, Czech Republik
Tomáš Lengyelfalusi, DTI, Slovakia
Karel Lepka, Masaryk University in Brno, Czech Republik
Šárka Mayerová, University of Defence, Brno, Czech Republik
Miroslava Růžičková, University of Bialystok, Poland
Andrey Shatyrko, Kiev National University, Ukraine
Zdeněk Svoboda, Brno University of Technology, Czech Republik
Zdeněk Šmarda, Brno University of Technology, Czech Republik

BOUNDED SOLUTIONS TO NONLINEAR TRIANGULAR SYSTEMS OF DISCRETE EQUATIONS

Jaromír Baštinec, Josef Diblík, Zuzana Piskořová
Brno University of Technology,
FEEC, Technická 10, 616 00 Brno, Czech Republic
bastinec@feec.vutbr.cz, diblik@feec.vutbr.cz, xpisko01@stud.feec.vutbr.cz

Abstract: *A nonlinear triangular system of discrete equations*

$$u_i(k+1) = q_i(k) \prod_{j=1}^{i} u_j^{p_{ij}}(k),\ i = 1, \dots, n,$$

is considered where $k \in \{a, a+1, \dots\}$, a is a fixed positive integer, q_i are real functions and exponents p_{ij} are positive constants. Sufficient conditions are formulated assuring the existence of at least one solution $u = u(k)$, $k \in \{a, a+1, \dots\}$ such that its coordinates $u_i(k)$, $i = 1, \dots, n$, are bounded above and below by given functions. Two convergent sequences of functions are constructed such that, with their limits, it is possible to define a set of initial values generating such solutions.

Keywords: nonlinear triangular system, discrete equation, convergent sequence.

INTRODUCTION

Denote by $\mathcal{N}(a)$ the set $\{a, a+1, \dots\}$, where a is a fixed positive integer. In the paper we study the following nonlinear triangular systems of discrete equations

$$u_1(k+1) = q_1(k) u_1^{p_{11}}(k), \tag{1}$$
$$u_2(k+1) = q_2(k) u_1^{p_{21}}(k) u_2^{p_{22}}(k), \tag{2}$$
$$u_3(k+1) = q_3(k) u_1^{p_{31}}(k) u_2^{p_{32}}(k) u_3^{p_{33}}(k), \tag{3}$$
$$\vdots$$
$$u_i(k+1) = q_i(k) u_1^{p_{i1}}(k) u_2^{p_{i2}}(k) u_3^{p_{i3}}(k) \dots u_i^{p_{ii}}(k), \tag{4}$$
$$\vdots$$
$$u_n(k+1) = q_n(k) u_1^{p_{n1}}(k) u_2^{p_{n2}}(k) u_3^{p_{n3}}(k) \dots u_i^{p_{ni}}(k) u_{i+1}^{p_{n,i+1}}(k) \dots u_n^{p_{nn}}(k) \tag{5}$$

where $u = (u_1, u_2, \dots, u_n)$ is a vector of unknown variables, $q_i \colon \mathcal{N}(a) \to (0, \infty)$, $i = 1, \dots, n$, are given functions and $p_{ij} \in (0, \infty)$, $i = 1, \dots, n$, $j = 1, 2, \dots, i$ are given powers of coordinates of unknown variables. By a solution of the system (1)–(5) we mean a vector $u = u^*(k) = (u_1^*(k), u_2^*(k), \dots, u_n^*(k))$ where $u_i^* \colon \mathcal{N}(a) \to \mathbb{R}$, $i = 1, \dots, n$ such that every equation of the system is satisfied for all $k \in \mathcal{N}(a)$ if $u_i(k)$ is replaced by $u_i^*(k)$, $i = 1, \dots, n$.

The system (1)–(5) can be written briefly as

$$u_i(k+1) = q_i(k) \prod_{j=1}^{i} u_j^{p_{ij}}(k), \;\; k \in \mathcal{N}(a), \; i = 1, \ldots, n. \tag{6}$$

In the paper sufficient conditions are found guaranteeing the existence of a nontrivial solution $u(k)$, $k \in \mathcal{N}(a)$ of the system (6) such that its coordinates $u_i(k)$, $i = 1, \ldots, n$, remain, for every $k \in \mathcal{N}(a)$, bounded above and below by given functions. Next, two types of monotone convergent sequences are constructed such that, with their limits, it is possible to define a set of initial values generating such solutions. The present paper generalizes some of previous results derived for the case of a scalar equation and for a system of two equations in [1,2,4] and continues with analysis of asymptotic behavior of solutions of discrete equations [3]. Throughout the paper we use following definition - when symbols for products are applied, we define $\prod_{j=p_1}^{p_2} \ldots := 1$ (where by dots a relevant argument is denoted) whenever, for integers p_1 and p_2, inequality $p_1 > p_2$ holds.

1 EXISTENCE OF SOLUTIONS WITH PRESCRIBED BEHAVIOR

Throughout the paper we assume that functions $b_i, c_i \colon \mathcal{N}(a) \to R$, $i = 1, \ldots, n$, are given and satisfy

$$0 \le b_i(k) < c_i(k), \quad \forall k \in \mathcal{N}(a), \;\; i = 1, \ldots, n. \tag{7}$$

The following theorem provides sufficient conditions for the existence of at least one solution $u(k) = (u_1(k), u_2(k), \ldots, u_n(k))$, $k \in \mathcal{N}(a)$, to system (6), $i = 1, \ldots, n$, such that its ith coordinate, $i = 1, \ldots, n$ satisfies $b_i(k) < u_i(k) < c_i(k)$, $k \in \mathcal{N}(a)$.

Theorem 1 *Let functions $b_i, c_i \colon \mathcal{N}(a) \to R$, $i = 1, 2, \ldots, n$ satisfy inequalities* (7). *Assume that, for every $k \in \mathcal{N}(a)$ and every $i = 1, \ldots, n$,*

$$q_i(k) \left(\prod_{j=1}^{i-1} c_j^{p_{ij}}(k) \right) b_i^{p_{ii}}(k) < b_i(k+1) \tag{8}$$

and

$$q_i(k) \left(\prod_{j=1}^{i-1} b_j^{p_{ij}}(k) \right) c_i^{p_{ii}}(k) > c_i(k+1). \tag{9}$$

Then, there exists a solution $u(k) = (u_1(k), u_2(k), \ldots, u_n(k))$, $k \in \mathcal{N}(a)$ to system (6) *satisfying*

$$b_i(k) < u_i(k) < c_i(k) \;\; \forall k \in \mathcal{N}(a), \; i = 1, 2, \ldots, n. \tag{10}$$

We omit the proof referring to methods suggested in [5, 6]. Using [5, Theorem 1] or [6, Theorem 2] the theorem can be proved.

2 SEQUENCES GENERATING INITIAL VALUES

First, we will construct special monotone and convergent sequences $\{u_{ics}\}_{s=0}^{\infty}$ and $\{u_{ibs}\}_{s=0}^{\infty}$, where $i = 1, \dots, n$,, such that, using their limits, sets can be defined of initial values generating solutions with the behaviour described in Theorem 1 by formula (10) and next, we formulate a test if only one such solution exists.

Let us explain how to construct the mentioned sequences $\{u_{ics}\}_{s=0}^{\infty}$ and $\{u_{ibs}\}_{s=0}^{\infty}$, $i = 1, \dots, n$,. Below we assume that the hypotheses of Theorem 1 hold.

2.1 CONSTRUCTION OF SEQUENCES $\{u_{ics}\}_{s=0}^{\infty}$ and $\{u_{ibs}\}_{s=0}^{\infty}$, $i = 1, \dots, n$

The property of the terms of the sequence $\{u_{ics}\}_{s=0}^{\infty}$, $i = 1, \dots, n$ is the following. For every fixed $s \in \mathcal{N}(0)$, the solution of the initial problem

$$(u_1(a), \dots, u_i(a)) = (u_{1cs}, \dots, u_{ics}), \;\; i = 1, \dots, n,$$

for system (1)–(5) defines a solution $u(k) = u_s(k) = (u_{1s}(k), \dots, u_{is}(k))$ such that

$$u_{is}(a+s) = c_i(a+s), \;\; i = 1, \dots, n. \tag{11}$$

Similarly, the terms of the sequence $\{u_{ibs}\}_{s=0}^{\infty}$, $i = 1, \dots, n$ are such that the initial problem

$$(u_1(a), \dots, u_i(a)) = (u_{1bs}, \dots, u_{ibs}), \;\; i = 1, \dots, n,$$

for system (1)–(5) defines a solution $u(k) = u_s(k) = (u_{1s}(k), \dots, u_{is}(k))$ such that

$$u_{is}(a+s) = b_i(a+s), \;\; i = 1, \dots, n. \tag{12}$$

2.1.1 CONSTRUCTION OF SEQUENCES $\{u_{1cs}\}_{s=0}^{\infty}$ and $\{u_{1bs}\}_{s=0}^{\infty}$

Now we will construct explicit formulas for the terms of sequences $\{u_{1cs}\}_{s=0}^{\infty}$, $\{u_{1cs}\}_{s=0}^{\infty}$. Consider equation (1). Then its general solution is expressed by the formula

$$u_1(a+s) = \left(\prod_{l=0}^{s-1} q_1^{p_{11}^{s-1-l}}(a+l) \right) u_1^{p_{11}^s}(a), \;\; s = 0, \dots \,. \tag{13}$$

Assuming, in accordance with (11), $u_{1s}(a+s) = c_1(a+s)$, we express $u_1(a)$ from (13), and define

$$u_{1cs} := u_1(a) = \left[c_1(a+s) \left(\prod_{l=0}^{s-1} q_1^{p_{11}^{s-1-l}}(a+l) \right)^{-1} \right]^{1/p_{11}^s}, \quad s = 0, \dots \,.$$

Similarly, expressing $u_1(a)$ from formula (13), assuming $u_{1s}(a+s) = b_1(a+s)$ by (12) , we define

$$u_{1bs} := u_1(a) = \left[b_1(a+s) \left(\prod_{l=0}^{s-1} q_1^{p_{11}^{s-1-l}}(a+l) \right)^{-1} \right]^{1/p_{11}^s}, \quad s = 0, \dots \,.$$

Let us prove that the sequence $\{u_{1cs}\}_{s=0}^{\infty}$ is decreasing and the sequence $\{u_{1bs}\}_{s=0}^{\infty}$ is increasing. Simplifying inequality

$$u_{1c,s+1} < u_{1cs}, \quad s = 0, \ldots, \tag{14}$$

that is, the inequality

$$\left[c_1(a+s+1) \left(\prod_{l=0}^{s} q_1^{p_{11}^{s-l}}(a+l) \right)^{-1} \right]^{1/p_{11}^{s+1}} < \left[c_1(a+s) \left(\prod_{l=0}^{s-1} q_1^{p_{11}^{s-1-l}}(a+l) \right)^{-1} \right]^{1/p_{11}^{s}},$$

we get

$$c_1(a+s+1) < c_1^{p_{11}}(a+s) q_1(a+s). \tag{15}$$

This inequality holds for every $s = 0, \ldots$, because it is equivalent with assumption (9) where $i = 1$. Due to properties of the functions $c_1(k)$, $q_1(k)$, $k = 0, \ldots$, and positivity of p_{11}, from inequality (15) follows inequality (14). Similarly, simplifying inequality

$$u_{1b,s+1} > u_{1bs}, \quad s = 0, \ldots,$$

that is, the inequality

$$\left[b_1(a+s+1) \left(\prod_{l=0}^{s} q_1^{p_{11}^{s-l}}(a+l) \right)^{-1} \right]^{1/p_{11}^{s+1}} > \left[b_1(a+s) \left(\prod_{l=0}^{s-1} q_1^{p_{11}^{s-1-l}}(a+l) \right)^{-1} \right]^{1/p_{11}^{s}},$$

we derive an equivalent inequality

$$b_1(a+s+1) > b_1^{p_{11}}(a+s) q_1(a+s).$$

The latter inequality holds because it is a variant of (8) where $i = 1$. Finally, let us show that

$$u_{1bs} < u_{1cs}, \quad s = 0, \ldots . \tag{16}$$

For $s = 0$, inequality (16) turns into

$$u_{1b0} = b_1(0) < c_1(0) = u_{1c0}$$

and is a consequence of (7) where $k = 0$. Let $s > 0$. Then the awaited inequality

$$\left[b_1(a+s) \left(\prod_{l=0}^{s-1} q_1^{p_{11}^{s-1-l}}(a+l) \right)^{-1} \right]^{1/p_{11}^{s}} < \left[c_1(a+s) \left(\prod_{l=0}^{s-1} q_1^{p_{11}^{s-1-l}}(a+l) \right)^{-1} \right]^{1/p_{11}^{s}}$$

immediately implies $b_1(a+s) < c_1(a+s)$, which is valid due to (7) and vice versa. Therefore, both sequences $\{u_{1cs}\}_{s=0}^{\infty}$ and $\{u_{1bs}\}_{s=0}^{\infty}$ are convergent. Denote their limits as

$$u_{1b}^* = \lim_{s \to \infty} u_{1bs}, \quad u_{1c}^* = \lim_{s \to \infty} u_{1cs}.$$

Obviously, the set

$$I_1 := [u_{1b}^*, u_{1c}^*]$$

is nonempty, since it contains at least one point (in this case $u_{1b}^* = u_{1c}^*$).

2.1.2 CONSTRUCTION OF SEQUENCES $\{u_{2cs}\}_{s=0}^{\infty}$ and $\{u_{2bs}\}_{s=0}^{\infty}$

Let $u_1^* \in I_1$ be a fixed point. Then the initial point

$$(a, u_1(a)) = (a, u_1^*)$$

defines a solution $u_1 = u_1^*(k)$, $k =$ of equation (1), see (13),

$$u_1^*(k) = \left(\prod_{l=0}^{k-a-1} q_1^{p_{11}^{k-a-1-l}}(a+l) \right) (u_1^*)^{p_{11}^s}, \quad k = a, \dots . \tag{17}$$

Consider equation (2) where $u_1(k) := u_1^*(k)$, that is, the equation

$$u_2(k+1) = q_2(k)(u_1^*(k))^{p_{21}}(k) u_2^{p_{22}}(k). \tag{18}$$

General solution of equation (18) is expressed by the formula (compare with the formula (13))

$$u_2(k) = \left(\prod_{l=0}^{k-a-1} [q_2(a+l)(u_1^*(a+l))^{p_{21}}]^{p_{22}^{k-a-1-l}} \right) u_2^{p_{22}^s}(a), \quad k = a, \dots . \tag{19}$$

Assuming, in accordance with (11), $u_{2s}(a+s) = c_2(a+s)$, we express $u_2(a)$ from (19), and define

$$u_{2cs} := u_2(a) = \left[c_2(a+s) \left(\prod_{l=0}^{s-1} [q_2(a+l)(u_1^*(a+l))^{p_{21}}]^{p_{22}^{s-1-l}} \right)^{-1} \right]^{1/p_{22}^s}, \quad s = 0, \dots .$$

Similarly, expressing $u_2(a)$ from formula (19), assuming $u_{2s}(a+s) = b_2(a+s)$ by (12), we define

$$u_{2bs} := u_2(a) = \left[b_2(a+s) \left(\prod_{l=0}^{s-1} [q_2(a+l)(u_1^*(a+l))^{p_{21}}]^{p_{22}^{s-1-l}} \right)^{-1} \right]^{1/p_{22}^s}, \quad s = 0, \dots .$$

Let us prove that the sequence $\{u_{2cs}\}_{s=0}^{\infty}$ is decreasing and the sequence $\{u_{2bs}\}_{s=0}^{\infty}$ is increasing. Simplifying inequality $u_{2c,s+1} < u_{2cs}$, $s = 0, \dots$, that is, the inequality

$$\left[c_2(a+s+1) \left(\prod_{l=0}^{s} [q_2(a+l)(u_1^*(a+l))^{p_{21}}]^{p_{22}^{s-l}} \right)^{-1} \right]^{1/p_{22}^{s+1}}$$

$$< \left[c_2(a+s) \left(\prod_{l=0}^{s-1} [q_2(a+l)(u_1^*(a+l))^{p_{21}}]^{p_{22}^{s-1-l}} \right)^{-1} \right]^{1/p_{22}^s}, \quad s = 0, \dots,$$

we get

$$c_2(a+s+1) < c_2^{p_{22}}(a+s) q_2(a+s)(u_1^*(a+s))^{p_{21}}. \tag{20}$$

This inequality holds for every $s = 0, \dots$, because solution $u_1^*(k)$, $k = a, a+1, \dots$, satisfies inequalities (10) where $i = 1$, the right-hand side of (20) can be estimated as

$$c_2^{p_{22}}(a+s) q_2(a+s)(u_1^*(a+s))^{p_{21}} > c_2^{p_{22}}(a+s) q_2(a+s)(b_1(a+s))^{p_{21}}$$

and (20) is valid due to assumption (9) with $i = 2$ (we do not mention other properties of $c_2(k)$, $q_2(k)$, $k = 0, \dots$, and p_{22}. Similarly, simplifying inequality

$$u_{2b,s+1} > u_{2bs}, \quad s = 0, \dots,$$

that is, the inequality

$$\left[b_2(a+s+1) \left(\prod_{l=0}^{s} [q_2(a+l)(u_1^*(a+l))^{p_{21}}]^{p_{22}^{s-l}} \right)^{-1} \right]^{1/p_{22}^{s+1}}$$

$$> \left[b_2(a+s) \left(\prod_{l=0}^{s-1} [q_2(a+l)(u_1^*(a+l))^{p_{21}}]^{p_{22}^{s-1-l}} \right)^{-1} \right]^{1/p_{22}^{s}},$$

we get

$$b_2(a+s+1) > b_2^{p_{22}}(a+s) q_2(a+s)(u_1^*(a+s))^{p_{21}}. \tag{21}$$

This inequality holds for every $s = 0, \dots$, because solution $u_1^*(k)$, $k = a, a+1, \dots$ satisfies inequalities (10) where $i = 1$, the right-hand side of (21) can be estimated as

$$b_2^{p_{22}}(a+s) q_2(a+s)(u_1^*(a+s))^{p_{21}} < b_2^{p_{22}}(a+s) q_2(a+s)(c_1(a+s))^{p_{21}}$$

and (21) is valid due to assumption (8) with $i = 2$ (we do not mention other properties of $c_2(k)$, $q_2(k)$, $k = 0, \dots$, and p_{22}). Finally, let us show that

$$u_{2bs} < u_{2cs}, \quad s = 0, \dots . \tag{22}$$

For $s = 0$, inequality (22) turns into

$$u_{2b0} = b_2(0) < c_2(0) = u_{2c0}$$

and is a consequence of (7) where $k = 0$. Let $s > 0$. Then the awaited inequality

$$\left[b_2(a+s) \left(\prod_{l=0}^{s-1} [q_2(a+l)(u_1^*(a+l))^{p_{21}}]^{p_{22}^{s-1-l}} \right)^{-1} \right]^{1/p_{22}^{s}}$$

$$< \left[c_2(a+s) \left(\prod_{l=0}^{s-1} [q_2(a+l)(u_1^*(a+l))^{p_{21}}]^{p_{22}^{s-1-l}} \right)^{-1} \right]^{1/p_{22}^{s}}$$

immediately implies $b_2(a+s) < c_2(a+s)$, which is valid due to (7) and vice versa. Therefore, both sequences $\{u_{2cs}\}_{s=0}^{\infty}$ and $\{u_{2bs}\}_{s=0}^{\infty}$ are convergent. Denote their limits as

$$u_{2b}^* = \lim_{s \to \infty} u_{2bs}, \quad u_{2c}^* = \lim_{s \to \infty} u_{2cs}.$$

Obviously, the set $I_2 := [u_{2b}^*, u_{2c}^*]$ is nonempty, since it contains at least one point (in this case $u_{2b}^* = u_{2c}^*$). Because we assumed, constructing the sequences $\{u_{2cs}\}_{s=0}^{\infty}$, $\{u_{2bs}\}_{s=0}^{\infty}$, that a solution $u_1^*(k)$, $k = a, \dots$ of equation (1), defined by the initial point $u_1^* \in I_1$, is fixed, the values of the above limits depend on u_1^* as well as interval I_2. Then we will below indicate this dependence writing $u_{2b}^* = u_{2b}^*(u_1^*)$, $u_{2c}^* = u_{2c}^*(u_1^*)$ and

$$I_2^* := I_2(u_1^*) = [u_{2b}^*(u_1^*), u_{2c}^*(u_1^*)].$$

2.1.3 CONSTRUCTION OF SEQUENCES $\{u_{3cs}\}_{s=0}^{\infty}$ and $\{u_{3bs}\}_{s=0}^{\infty}$

Let points $u_1^* \in I_1$ and $u_2^* \in I_2^*$ be fixed. These two points define, by formula (17), solution $u_1^*(k)$, $k = a, \dots$ of equation (1) satisfying inequality (10) with $i = 1$ and solution of equation (2)

$$u_2^*(k) = \left(\prod_{l=0}^{k-a-1} [q_2(a+l)(u_1^*(a+l))^{p_{21}}]^{p_{22}^{k-a-1-l}} \right) (u_2^*)^{p_{22}^s}(a), \quad k = a, \dots, \tag{23}$$

satisfying inequality (10) with $i = 2$. Consider equation (3) where $u_i(k) := u_i^*(k)$, $i = 1, 2$, that is, the equation

$$u_3(k+1) = q_3(k)(u_1^*(k))^{p_{31}}(u_2^*(k))^{p_{32}} u_3^{p_{33}}(k). \tag{24}$$

General solution of equation (24) is expressed by the formula (compare with formulas (19), (23))

$$u_3(k) = \left(\prod_{l=0}^{k-a-1} [q_3(a+l)(u_1^*(a+l))^{p_{31}}(u_2^*(a+l))^{p_{32}}]^{p_{33}^{k-a-1-l}} \right) u_3^{p_{33}^s}(a), \quad k = a, \dots . \tag{25}$$

Assuming, in accordance with (11), $u_{s3}(a+s) = c_3(a+s)$, we express $u_3(a)$ from (25), and define

$$u_{3cs} := u_3(a) = \left[c_3(a+s) \left(\prod_{l=0}^{s-1} [q_3(a+l)(u_1^*(a+l))^{p_{31}}(u_2^*(a+l))^{p_{32}}]^{p_{33}^{s-1-l}} \right)^{-1} \right]^{1/p_{33}^s}$$

where $s = 0, \dots$. Similarly, expressing $u_3(a)$ from formula (25), assuming $u_{3s}(a+s) = b_3(a+s)$ by (12), we derive

$$u_{3bs} := u_3(a) = \left[b_3(a+s) \left(\prod_{l=0}^{s-1} [q_3(a+l)(u_1^*(a+l))^{p_{31}}(u_2^*(a+l))^{p_{32}}]^{p_{33}^{s-1-l}} \right)^{-1} \right]^{1/p_{33}^s}$$

where $s = 0, \dots$. Let us prove that the sequence $\{u_{3cs}\}_{s=0}^{\infty}$ is decreasing and the sequence $\{u_{3bs}\}_{s=0}^{\infty}$ is increasing. Simplifying inequality

$$u_{3c,s+1} < u_{3cs}, \quad s = 0, \dots, \tag{26}$$

that is, the inequality

$$\left[c_3(a+s+1) \left(\prod_{l=0}^{s} [q_3(a+l)(u_1^*(a+l))^{p_{31}}(u_2^*(a+l))^{p_{32}}]^{p_{33}^{s-l}} \right)^{-1} \right]^{1/p_{33}^{s+1}}$$
$$< \left[c_3(a+s) \left(\prod_{l=0}^{s-1} [q_3(a+l)(u_1^*(a+l))^{p_{31}}(u_2^*(a+l))^{p_{32}}]^{p_{33}^{s-1-l}} \right)^{-1} \right]^{1/p_{33}^s}, \quad s = 0, \dots .$$

We get

$$c_3(a+s+1) < c_3^{p_{33}}(a+s) q_3(a+s)(u_1^*(a+s))^{p_{31}}(u_2^*(a+s))^{p_{32}}. \tag{27}$$

This inequality holds for every $s = 0, \dots$, because solution $u_1^*(k)$, $k = a, a+1, \dots$, satisfies inequalities (10) where $i = 1$, solution $u_2^*(k)$, $k = a, a+1, \dots$, satisfies inequalities (10) where $i = 2$, the right-hand side of (27) can be estimated as

$$c_3^{p_{33}}(a+s)q_3(a+s)(u_1^*(a+s))^{p_{31}}(u_2^*(a+s))^{p_{32}} \\ > c_3^{p_{33}}(a+s)q_3(a+s)(b_1(a+s))^{p_{31}}(b_2(a+s))^{p_{32}}$$

and (27) is valid due to assumption (9) with $i = 3$ (we do not mention other properties of $c_3(k)$, $q_3(k)$, $k = 0, \dots$, and p_{33}). Similarly, simplifying inequality

$$u_{3b,s+1} > u_{3bs}, \quad s = 0, \dots,$$

that is, the inequality

$$\left[b_3(a+s+1) \left(\prod_{l=0}^{s} [q_3(a+l)(u_1^*(a+l))^{p_{31}}(u_2^*(a+l))^{p_{32}}]^{p_{33}^{s-l}} \right)^{-1} \right]^{1/p_{33}^{s+1}} \\ > \left[b_3(a+s) \left(\prod_{l=0}^{s-1} [q_3(a+l)(u_1^*(a+l))^{p_{31}}(u_2^*(a+l))^{p_{32}}]^{p_{33}^{s-1-l}} \right)^{-1} \right]^{1/p_{33}^{s}}, \quad s = 0, \dots,$$

we get

$$b_3(a+s+1) > b_3^{p_{33}}(a+s)q_3(a+s)(u_1^*(a+s))^{p_{31}}(u_2^*(a+s))^{p_{32}}. \tag{28}$$

This inequality holds for every $s = 0, \dots$, because solution $u_1^*(k)$, $k = a, a+1, \dots$ satisfies inequalities (10) where $i = 1$, solution $u_2^*(k)$, $k = a, a+1, \dots$ satisfies inequalities (10) where $i = 2$, the right-hand side of (28) can be estimated as

$$b_3^{p_{33}}(a+s)q_3(a+s)(u_1^*(a+s))^{p_{31}}(u_2^*(a+s))^{p_{32}} \\ < b_3^{p_{33}}(a+s)q_3(a+s)(c_1(a+s))^{p_{31}}(c_2(a+s))^{p_{32}}$$

and (28) is valid due to assumption (8) with $i = 3$ (we do not mention other properties of $c_3(k)$, $q_3(k)$, $k = 0, \dots$, and p_{33}).

Finally, let us show that

$$u_{3bs} < u_{3cs}, \quad s = 0, \dots . \tag{29}$$

For $s = 0$, inequality (29) turns into

$$u_{3b0} = b_3(0) < c_3(0) = u_{3c0}$$

and it is a consequence of (7) where $k = 0$. Let $s > 0$. Then, the awaited inequality

$$\left[b_3(a+s) \left(\prod_{l=0}^{s-1} [q_3(a+l)(u_1^*(a+l))^{p_{31}}(u_2^*(a+l))^{p_{32}}]^{p_{33}^{s-1-l}} \right)^{-1} \right]^{1/p_{33}^{s}} \\ < \left[c_3(a+s) \left(\prod_{l=0}^{s-1} [q_3(a+l)(u_1^*(a+l))^{p_{31}}(u_2^*(a+l))^{p_{32}}]^{p_{33}^{s-1-l}} \right)^{-1} \right]^{1/p_{33}^{s}}$$

immediately implies $b_3(a+s) < c_3(a+s)$ which is valid due to (7) and vice versa. Therefore, both sequences $\{u_{3cs}\}_{s=0}^{\infty}$ and $\{u_{3bs}\}_{s=0}^{\infty}$ are convergent. Denote their limits as

$$u_{3b}^* = \lim_{s\to\infty} u_{3bs}, \quad u_{3c}^* = \lim_{s\to\infty} u_{3cs}.$$

Obviously, the set $I_3 := [u_{3b}^*, u_{3c}^*]$ is nonempty, since it contains at least one point (in this case $u_{3b}^* = u_{3c}^*$).

Because we assumed, constructing the sequences $\{u_{3cs}\}_{s=0}^{\infty}$, $\{u_{3bs}\}_{s=0}^{\infty}$, that a solution $u_1^*(k)$, $k = a, \dots$, of equation (1), defined by the initial point $u_1^* \in I_1$, is fixed and a solution $u_2^*(k)$, $k = a, \dots$, of equation (2), defined by the initial point $u_2^* \in I_2^*$, is fixed, the values of the above limits depend on u_1^*, u_2^* as well as the interval I_3. Then, we will below indicate this dependence and we will write

$$u_{3b}^* = u_{3b}^*(u_1^*, u_2^*), \quad u_{3c}^* = u_{3c}^*(u_1^*, u_2^*)$$

and

$$I_3^* := I_3(u_1^*, u_2^*) = [u_{3b}^*(u_1^*, u_2^*), u_{3c}^*(u_1^*, u_2^*)].$$

2.1.4 CONSTRUCTION OF SEQUENCES $\{u_{ics}\}_{s=0}^{\infty}$ and $\{u_{ibs}\}_{s=0}^{\infty}$, $i = 4, \dots, n$

Let $n \geq 4$ and let an index $j \in \{3, n-1\}$ be fixed. By induction, assume that solutions $u_i^*(k)$, $k = a, \dots,\ i = 1, \dots, j$, of equations (1)–(4) are constructed by the above formulated approach and satisfy inequalities (10) where $i = 1, \dots, j$. Assume as well that these solutions are determined by initial values $u_i^*(a) = u_i^*$, $i = 1, \dots, j$, where $u_i^* \in I_i^*$, $i = 1, \dots, j$, $I_1^* := I_1$ and

$$I_s^* = I_s(u_1^*, u_2^*, \dots, u_{s-1}^*) = [u_{sb}^*(u_1^*, u_2^* \dots, u_{s-1}^*), u_{sc}^*(u_1^*, u_2^* \dots, u_{s-1}^*)], \quad s = 2, \dots, j.$$

Consider equation (4) where $i = j+1$ and $u_i = u_i^*(k)$, $k = a, \dots$, $i = 1, \dots, j$ are solutions of equations (1)–(4) mentioned above, that is the equation

$$u_{j+1}(k+1) = q_{j+1}(k)(u_1^*(k))^{p_{j+1,1}} \dots (u_j^*(k))^{p_{j+1,j}} u_{j+1}^{p_{j+1,j+1}}(k). \tag{30}$$

General solution of equation (30) is expressed by the formula (compare with formulas (19), (23) (25))

$$u_{j+1}(k) = \left(\prod_{l=0}^{k-a-1} [q_{j+1}(a+l)(u_1^*(a+l))^{p_{j+1,1}} \dots (u_j^*(a+l))^{p_{j+1,j}}]^{p_{j+1,j+1}^{k-a-1-l}} \right) u_{j+1}^{p_{j+1,j+1}^s}(a) \tag{31}$$

where $k = a, \dots$. Assuming, in accordance with (11), $u_{s,j+1}(a+s) = c_{j+1}(a+s)$, we express $u_{j+1}(a)$ from (31), and define

$$\begin{aligned} u_{j+1,cs} &:= u_{j+1}(a) \\ &= \left[c_{j+1}(a+s) \left(\prod_{l=0}^{s-1} [q_{j+1}(a+l)(u_1^*(a+l))^{p_{j+1,1}} \dots (u_j^*(a+l))^{p_{j+1,j}}]^{p_{j+1,j+1}^{s-1-l}} \right)^{-1} \right]^{1/p_{j+1,j+1}^s} \end{aligned}$$

where $s = 0, \dots$. Similarly, expressing $u_{j+1}(a)$ from formula (31), assuming $u_{j+1,s}(a+s) = b_{j+1}(a+s)$ by (12), we derive

$$u_{j+1,bs} := u_{j+1}(a)$$
$$= \left[b_{j+1}(a+s) \left(\prod_{l=0}^{s-1} [q_{j+1}(a+l)(u_1^*(a+l))^{p_{j+1,1}} \dots (u_j^*(a+l))^{p_{j+1,j}}]^{p_{j+1,j+1}^{s-1-l}} \right)^{-1} \right]^{1/p_{j+1,j+1}^s}$$

where $s = 0, \dots$. Let us prove that the sequence $\{u_{j+1,cs}\}_{s=0}^{\infty}$ is decreasing and the sequence $\{u_{j+1,bs}\}_{s=0}^{\infty}$ is increasing. Simplifying inequality $u_{j+1,c,s+1} < u_{j+1,cs}$, $s = 0, \dots$, that is, the inequality

$$\left[c_{j+1}(a+s+1) \left(\prod_{l=0}^{s} [q_{j+1}(a+l)(u_1^*(a+l))^{p_{j+1,1}} \dots (u_j^*(a+l))^{p_{j+1,j}}]^{p_{j+1,j+1}^{s-l}} \right)^{-1} \right]^{1/p_{j+1,j+1}^{s+1}}$$
$$< \left[c_{j+1}(a+s) \left(\prod_{l=0}^{s-1} [q_{j+1}(a+l)(u_1^*(a+l))^{p_{j+1,1}} \dots (u_j^*(a+l))^{p_{j+1,j}}]^{p_{j+1,j+1}^{s-1-l}} \right)^{-1} \right]^{1/p_{j+1,j+1}^{s}}$$

where $s = 0, \dots$, we get

$$c_{j+1}(a+s+1) < c_{j+1}^{p_{j+1,j+1}}(a+s) q_{j+1}(a+s)(u_1^*(a+s))^{p_{j+1,1}} \dots (u_j^*(a+s))^{p_{j+1,j}}. \qquad (32)$$

We show that this inequality holds for every $s = 0, \dots$, because solution $u_1^*(k)$, $k = a, a+1, \dots$, satisfies inequalities (10) where $i = 1$, solution $u_2^*(k)$, $k = a, a+1, \dots$, satisfies inequalities (10), where $i = 2$, etc. and solution $u_j^*(k)$, $k = a, a+1, \dots$, satisfies inequalities (10) where $i = j$. Then, the right-hand side of (32) can be estimated as

$$c_{j+1}^{p_{j+1,j+1}}(a+s) q_{j+1}(a+s)(u_1^*(a+s))^{p_{j+1,1}} \dots (u_j^*(a+s))^{p_{j+1,j}}$$
$$> c_{j+1}^{p_{j+1,j+1}}(a+s) q_{j+1}(a+s)(b_1(a+s))^{p_{j+1,1}} \dots (b_j(a+s))^{p_{j+1,j}}$$

and (32) is valid due to assumption (9) with $i = j+1$ (we do not mention other properties of $c_{j+1}(k)$, $q_{j+1}(k)$, $k = 0, \dots$, and $p_{j+1,j+1}$). Similarly, simplifying inequality

$$u_{j+1,b,s+1} > u_{jbs}, \quad s = 0, \dots,$$

that is, the inequality

$$\left[b_{j+1}(a+s+1) \left(\prod_{l=0}^{s} [q_{j+1}(a+l)(u_1^*(a+l))^{p_{j+1,1}} \dots (u_j^*(a+l))^{p_{j+1,j}}]^{p_{j+1,j+1}^{s-l}} \right)^{-1} \right]^{1/p_{j+1,j+1}^{s+1}}$$
$$> \left[b_{j+1}(a+s) \left(\prod_{l=0}^{s-1} [q_{j+1}(a+l)(u_1^*(a+l))^{p_{j+1,1}} \dots (u_j^*(a+l))^{p_{j+1,j}}]^{p_{j+1,j+1}^{s-1-l}} \right)^{-1} \right]^{1/p_{j+1,j+1}^{s}}$$

where $s = 0, \ldots$, we get

$$b_{j+1}(a+s+1) > b_{j+1}^{p_{j+1,j+1}}(a+s)q_{j+1}(a+s)(u_1^*(a+s))^{p_{j+1,1}} \ldots (u_j^*(a+s))^{p_{j+1,j}}. \tag{33}$$

This inequality holds for every $s = 0, \ldots$, because solution $u_1^*(k)$, $k = a, a+1, \ldots$, satisfies inequalities (10) where $i = 1$, solution $u_2^*(k)$, $k = a, a+1, \ldots$, satisfies inequalities (10) where $i = 2$, etc. and solution $u_j^*(k)$, $k = a, a+1, \ldots$, satisfies inequalities (10) where $i = j$. The right-hand side of (33) can be estimated as

$$b_{j+1}^{p_{j+1,j+1}}(a+s)q_{j+1}(a+s)(u_1^*(a+s))^{p_{j+1,1}} \ldots (u_j^*(a+s))^{p_{j+1,j}}$$
$$< b_{j+1}^{p_{j+1,j+1}}(a+s)q_{j+1}(a+s)(c_1(a+s))^{p_{j+1,1}} \ldots (c_j(a+s))^{p_{j+1,j}}$$

and (33) is valid due to assumption (8) with $i = j+1$ (we do not mention other properties of $b_{j+1}(k)$, $q_{j+1}(k)$, $k = 0, \ldots$, and $p_{j+1,j+1}$). Finally, let us show that

$$u_{j+1,bs} < u_{j+1,cs}, \quad s = 0, \ldots . \tag{34}$$

For $s = 0$, inequality (34) turns into

$$u_{j+1,b0} = b_{j+1}(0) < c_{j+1}(0) = u_{j+1,c0}$$

being a consequence of (7) where $k = 0$. Let $s > 0$. Then, the awaited inequality

$$\left[b_{j+1}(a+s) \left(\prod_{l=0}^{s-1} [q_{j+1}(a+l)(u_1^*(a+l))^{p_{j+1,1}} \ldots (u_j^*(a+l))^{p_{j+1,j}}]^{p_{j+1,j+1}^{s-1-l}} \right)^{-1} \right]^{1/p_{j+1,j+1}^s}$$
$$< \left[c_{j+1}(a+s) \left(\prod_{l=0}^{s-1} [q_{j+1}(a+l)(u_1^*(a+l))^{p_{j+1,1}} \ldots (u_j^*(a+l))^{p_{j+1,j}}]^{p_{j+1,j+1}^{s-1-l}} \right)^{-1} \right]^{1/p_{j+1,j+1}^s}$$

immediately implies $b_{j+1}(a+s) < c_{j+1}(a+s)$, which is valid due to (7) and vice versa. Therefore, both sequences $\{u_{j+1,cs}\}_{s=0}^{\infty}$ and $\{u_{j+1,bs}\}_{s=0}^{\infty}$ are convergent. Denote their limits as

$$u_{j+1,b}^* = \lim_{s\to\infty} u_{j+1,bs}, \quad u_{j+1,c}^* = \lim_{s\to\infty} u_{j+1,cs}.$$

Obviously, the set $I_{j+1} := [u_{j+1,b}^*, u_{j+1,c}^*]$ is nonempty, since it contains at least one point (in this case $u_{j+1,b}^* = u_{j+1c}^*$).

Because we assumed, constructing the sequences $\{u_{j+1,cs}\}_{s=0}^{\infty}$, $\{u_{j+1,bs}\}_{s=0}^{\infty}$, that a solution $u_1^*(k)$, $k = a, \ldots$, of equation (1), defined by the initial point $u_1^* \in I_1$, is fixed and a solution $u_2^*(k)$, $k = a, \ldots$, of equation (2), defined by the initial point $u_2^* \in I_2^*$, etc. and solution $u_j^*(k)$, $k = a, a+1, \ldots$, of equation (4) where $j = i$, defined by the initial point $u_j^* \in I_j$, is fixed, the values of above limits depend on $u_1^*, u_2^*, \ldots, u_{j^*}$, as well as the interval I_{j+1}. Then we will below indicate this dependence writing

$$u_{j+1,b}^* = u_{j+1,b}^*(u_1^*, u_2^*, \ldots, u_{j^*}), \quad u_{j+1,c}^* = u_{j+1,c}^*(u_1^*, u_2^*, \ldots, u_{j^*}) \tag{35}$$

and

$$I_{j+1}^* := I_{j+1}(u_1^*, u_2^*, , \ldots, u_{j^*}) = [u_{j+1,b}^*(u_1^*, u_2^*, \ldots, u_{j^*}), u_{j+1,c}^*(u_1^*, u_2^*, \ldots, u_{j^*})]. \tag{36}$$

The process of constructing the sequences will be terminated if $j = n-1$. Note that formulas (35), (36) hold (due to constructions in parts 2.1.1–2.1.4) for every $j = 0, \ldots, n-1$.

2.2 FINAL RESULTS

The proof of the following Theorem 2 and Theorem 3 follow directly from the properties of sequences constructed in parts 2.1.1–2.1.4. Therefore, we omit them.

Theorem 2 *Let the hypotheses of Theorem 1 hold. Then, every initial point* $(a, u_1^*, u_2^*, \ldots, u_n^*)$ *where*

$$u_i^* \in I_i^*, \;\; i = 1, 2, \ldots, n$$

and intervals I_i^* *are computed by formula* (36) *where* $j = 0, \ldots, n-1$*, defines a solution of the system* (6) *satisfying inequalities* (10)*.*

Theorem 3 *Let the hypotheses of Theorem 1 hold. If, moreover,*

$$\lim_{s\to\infty} \frac{(c_j(a+s))^{1/p_{jj}^s} - (b_j(a+s))^{1/p_{jj}^s}}{\left(\prod_{l=0}^{s-1}[q_j(a+l)(b_1(a+l))^{p_{j,1}} \ldots (b_{j-1}(a+l))^{p_{j,j-1}}]^{p_{jj}^{s-1-l}}\right)^{1/p_{jj}^s}} = 0,$$

$j = 1, \ldots, n$*, then there exists a unique solution of the system* (6)*, satisfying inequalities* (10)*.*

3 EXAMPLE

Let $n = 3$ and let a particular case of system (6)

$$u_1(k+1) = \frac{k^2}{k+1} u_1^2(k), \tag{37}$$

$$u_2(k+1) = \frac{k^4}{k+1} u_1^2(k) u_2^2(k), \tag{38}$$

$$u_3(k+1) = \frac{k^6}{k+1} u_1^2(k) u_2^2(k) u_3^2(k) \tag{39}$$

be specified where $q_i(k) = k^{2i}/(k+1)$, $i = 1, 2, 3$, $a = 1$ and $p_{ij} = 2$, $i, j = 1, 2, 3$, $j \le i$. Set

$$b_1(k) = \frac{1}{2k}, \quad b_2(k) = \frac{1}{5k}, \quad b_3(k) = \frac{1}{200k}$$

and

$$c_1(k) = \frac{2}{k}, \quad c_2(k) = \frac{5}{k}, \quad c_3(k) = \frac{200}{k}.$$

Let us verify inequalities (8) and (9). For $i = 1$ we have

$$q_1(k) b_1^2(k) = \frac{k^2}{k+1} \cdot \frac{1}{4k^2} = \frac{1}{4(k+1)} < \frac{1}{2(k+1)} = b_1(k+1)$$

and

$$q_1(k) c_1^2(k) = \frac{k^2}{k+1} \cdot \frac{4}{k^2} = \frac{4}{(k+1)} > \frac{2}{(k+1)} = c_1(k+1).$$

For $i = 2$ we obtain

$$q_2(k)c_1^2(k)b_2^2(k) = \frac{k^4}{k+1} \cdot \frac{4}{k^2} \cdot \frac{1}{25k^2} = \frac{4}{25(k+1)} < \frac{1}{5(k+1)} = b_2(k+1)$$

and

$$q_2(k)b_1^2(k)c_2^2(k) = \frac{k^4}{k+1} \cdot \frac{1}{4k^2} \cdot \frac{25}{k^2} = \frac{25}{4(k+1)} > \frac{5}{(k+1)} = c_2(k+1).$$

Finally, for $i = 3$, we derive

$$q_3(k)c_1^2(k)c_2^2(k)b_3^2(k) = \frac{k^6}{k+1} \cdot \frac{4}{k^2} \cdot \frac{25}{k^2} \cdot \frac{1}{40000k^2} = \frac{1}{400(k+1)} < \frac{1}{200(k+1)} = b_3(k+1)$$

and

$$q_3(k)b_1^2(k)b_2^2(k)c_3^2(k) = \frac{k^6}{k+1} \cdot \frac{1}{4k^2} \cdot \frac{1}{25k^2} \cdot \frac{40000}{k^2} = \frac{400}{k+1} > \frac{200}{(k+1)} = c_3(k+1).$$

All assumptions of Theorem 1 are fulfilled. Therefore the system (37)–(39) has a nontrivial solution $u = u^*(k) = (u_1^*(k), u_2^*(k), u_3^*(k))$ where $u_i^* \colon \mathcal{N}(1) \to \mathbb{R}$, $i = 1, 2, 3$ and

$$\frac{1}{2k} < u_1^*(k) < \frac{2}{k},$$

$$\frac{1}{5k} < u_2^*(k) < \frac{5}{k},$$

$$\frac{1}{200k} < u_3^*(k) < \frac{200}{k}.$$

CONCLUSIONS

In this paper we study the nonlinear triangular system (6). Conditions guaranted existence of a solution bounded from below and from above are formulated in Theorem 1. Moreover, Theorem 2 indicates how the initial values determinig such solutions can be found. Theorem 3 brings sufficient conditions for the existence only one such a solution. It is an open problem if results of the paper can be enlarged to more general nonlinear systems, for example, to a system

$$u_i(k+1) = q_i(k) \prod_{j=1}^{n} u_j^{p_{ij}}(k), \quad k \in \mathcal{N}(a), \ i = 1, \ldots, n$$

being more general than the system (6).

References

[1] J. Baštinec, J. Diblík, Initial data generating solutions of a power equation with certain asymptotic properties *3rd International Conference APLIMAT, Bratislava, STU, 2004*, 247–252.

[2] J. Baštinec, J. Diblík, Determination of initial data generating solutions of Bernoulli's type difference equations with prescribed asymptotic behavior, *Proceedings of the Eighth International Conference on Difference Equations and Applications* (Chapman & Hall/CRC, Boca Raton, FL, 2005), 39–49.

[3] J. Baštinec, J. Diblík, Two classes of positive solutions of a discrete equation, *International Conference on Mathematics, Information Technologies and Applied Sciences (MITAV) 2017*, Brno, 21-32,

[4] J. Baštinec, J. Diblík, E. Korobko, Bounded solutions of a triangular system of two nonlinear discrete equations, *Proceedings of the 18th International Conference of Numerical Analysis and Applied Mathematics, 2020, Rhodes, Greece*. In print.

[5] J. Diblík, Discrete retract principle for systems of discrete equations, *Comput. Math. Appl.* **42** (2001), 515–528.

[6] J. Diblík, Asymptotic behaviour of solutions of discrete equations, *Funct. Differ. Equ.* **11** (2004), 37–48.

Acknowledgement

This work has been supported by the grant of Faculty of Electrical Engineering and Communication, Brno University of Technology (research project No. FEKT-S-20-6225).

SPIRALS IN MATHEMATICS TEACHING

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

Abstract: *The article contains a few topics for motivation in the Mathematics teaching which are closely related to the possibilities of students´ research approach. Firstly, there are described several examples of mathematical spirals including historical remarks, further there are given these spirals´ equations in polar coordinates and in parametric formulations as well. In the conclusion there are derived the formulas for their lengths with the help of the integral calculus.*

Keywords: Length of curve, definite integral, spiral of Archimedes, logarithmic spiral, hyperbolic spiral, Fermat spiral.

INTRODUCTION

The notion "spiral" is used quite often in various meanings in the common life. However, the general public or students not concerned with mathematics could be surprised that the notion "spiral" appears in mathematics as well. There are series of mathematical spirals some of which will be dealt with in this article. At technical universities, this topic is discussed in detail within technical curves. Nevertheless, our students, future Mathematics teachers, hardly encounter the topic of spirals and technical curves during their study and thus this issue could play a motivational role while teaching Mathematics. Let us note that a precise formal description of mathematical curves is rather complicated and often requires practical knowledge of polar or parametric coordinates. Therefore, the study of spirals could help to introduce these coordinates to students and further encourage their deeper interest in mathematical analysis and differential geometry. Let us remind some necessary theoretical knowledge (See [1], [4], [7]).

Let $y = f(x)$ be a continuous real function on the interval $\langle a, b \rangle$. Then the graph of this function is a curve. The length of this curve is defined by an integral $\int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$.

If this curve is defined in polar coordinates by an equation $r = f(\varphi)$ pro $\alpha \leq \varphi \leq \beta$, then its length is determined by an integral $\int_\alpha^\beta \sqrt{r^2 + \left(\frac{dr}{d\varphi}\right)^2}\, d\varphi$.

If the curve is defined parametrically, i.e. $x = \varphi(t)$, $y = \psi(t)$, where t is a real number from the interval $\langle a, b \rangle$, the length of this curve is determined by an integral $\int_a^b \sqrt{\left(\frac{d\varphi}{dt}\right)^2 + \left(\frac{d\psi}{dt}\right)^2}\, dt$.

Generally, we can say that spirals are plane curves created by a point which performs the given movement on a line which is rotating evenly around its fixed point.

1. ARCHIMEDEAN SPIRAL

In his treatise "On spirals" (about 225 BC) Archimedes describes the origin of this curve approximately as follows ([12], pg. 154): "If a straight line of which one extremity remains fixed be made to revolve at a uniform rate in a plane until it returns to the position from which it started, and if, at the same time as the straight line revolves, a point moves at a uniform rate along the straight line, starting from the fixed extremity, the point will describe a spiral in the plane." This definition gives the oldest example when the curve is created as the result of a dual movement of a point (See Fig. 1, taken from [9]).

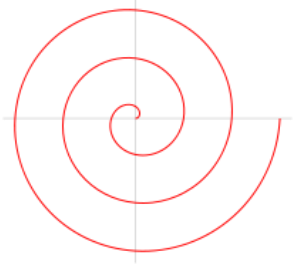


Fig. 1: Archimedean spiral
Source: [9]

Archimedean spirals can commonly be seen around us ([8]), for example in compressed springs, side edgings of rolled up carpets, on the rolled up rope or decorative spirals of jewellery. Among technical usages of Archimedean spirals, we can mention the transformation of the rotating movement to the linear one at sewing machines. Archimedean spiral appears at various mechanisms in machinery such as Archimedean screw. Based on this principle are constructed drills and screws. From the mathematical point of view, the most precise definition of Archimedean spiral is as a plane curve whose radius grows linearly with the angle. It can also be described as a trajectory of a point which moves equally along a half-line from its origin in point O, while the half-line rotates equally around the point O ([9]). The point O is a pole or the origin of a spiral. It is possible to prove that the ray coming from the origin of the spiral cuts the spiral in points whose distances from the pole form an arithmetic sequence.
In polar coordinates Archimedean spiral can be represented by an equation ([1], [5])

$$r = a\,\varphi\,, \text{ where } a \in \boldsymbol{R},\ a > 0,\ \varphi \in \boldsymbol{R},\ \varphi \geq 0.$$

It is more convenient to use the parametrical representation. For every point $X = [x, y]$ of the spiral there applies

$$x = r\cos t = at\cos t,\ y = r\sin t = at\sin t, \text{ where } t \text{ is a parameter, } t \in \boldsymbol{R},\ t > 0.$$

The length of Archimedean spiral is finite and we can calculate it using the formula for the length of a curve in polar coordinates. We will calculate the length s of a spiral from the origin O = [0, 0] to a point P of a spiral whose location is given by polar coordinates r, α. There applies $\frac{dr}{d\varphi} = a$, after modification

$$\sqrt{r^2 + \left(\frac{dr}{d\varphi}\right)^2} = \sqrt{r^2 + a^2} = a\sqrt{1+\varphi^2}\ .$$

We get

$$s = \int_0^{\alpha} a\sqrt{1+\varphi^2}\,d\varphi = \frac{a}{2}\left[\varphi\sqrt{1+\varphi^2} + ln\left(\varphi+\sqrt{1+\varphi^2}\right)\right]_0^{\alpha} = \frac{a}{2}\left[\alpha\sqrt{1+\alpha^2} + ln\left(\alpha+\sqrt{1+\alpha^2}\right)\right] .$$

This integral is quite difficult and while solving it students have to be patient and concentrated. Let $P = [r, \varphi]$, $Q = [s, \psi]$, $\varphi > \psi$, be two points of Archimedean spiral. It is possible to derive that for the area S of a sector POQ there applies $S = \frac{a^2}{6}\left(\varphi^3 - \psi^3\right)$. In conclusion, let us note that all values of angles in the above relations have to be expressed in radian measures. The given formulas including details can be found in [1].

2. LOGARITHMIC SPIRAL

Logarithmic spiral can be depicted verbally by several equivalent ways. In [11] there is given that a logarithmic spiral is a plane curve whose radius grows exponentially with the angle. In the book [1] we can find a description of a logarithmic spiral as a curve which intersects all half-lines leading from its origin with a constant angle α. There are two important points in this spiral: the pole and the origin of this spiral ([11]). The pole is the point around which the spiral "weaves". For the spiral in the basic shape (without translation) it is the point $[0, 0]$, i.e. the origin of the coordinate system. The origin of the spiral is the point from which the spiral is drawn. In the basic shape it is the point $[a, 0]$. The parameter a comes from mathematical representation of the logarithmic spiral ([11]). In the polar system of coordinates the logarithmic spiral can be represented by an equation

$$r = a\, e^{\,b\varphi},$$

or in the equivalent form by an equation

$$\varphi = \frac{1}{b} ln\left(\frac{r}{a}\right) ,$$

where a, b are positive real numbers and e is Euler's number. The ray coming out from the pole intersects the spiral in points whose distance from the pole create a geometric progression. The line joining the pole of the spiral with any of its point intersects the logarithmic spiral always under the same angle. Therefore, the logarithmic spiral is also called an equiangular spiral (René Descartes. 1638). The following picture shows an example of the logarithmic spiral (See [11]).

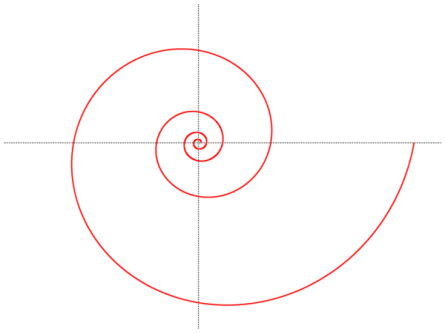


Fig. 2: Logarithmic spiral
Source: [11]

Let us note where we can meet this spiral in real life (See [1], [5] and [11]). The logarithmic spiral is used in the technical practice (rotating knives, cogwheels etc.) and at seamanship in

connection with loxodrome which is the most convenient curve for sailing and whose projection to the stereographic net is the logarithmic spiral. They also appear quite often in the nature; more precisely, in some natural phenomena there appear formations which resemble logarithmic spirals. Let us give some examples (taken from [5] and [11]):

- The trajectory on which birds of prey (hawks) approach their prey.The equiangularity of this spiral allows them to observe the pray under the constant angle.
- The trajectory on which insects approach the source of light.
- The arms of spiral galaxies. Galaxy Milky Way has several spiral arms and each arm corresponds roughly to a logarithmic spiral with an approximite angle of 12 degrees.
- Cloud belts created in the centres of tropical cyclones.
- A range of biological formations, e.g. mollusk shells, horns, elephant tusks, and spider webs.
- The arrangement of sunflower seeds.
- Double helical spiral of DNA.

The first person to mention a logarithmic spiral was a French mathematician René Descartes. Independently on Descartes, the logarithmic spiral was studied by Evangelist Torricelli, who set the formula for the length of the curve. Later, the spiral was studied thoroughly by Jacob Bernoulli. In this context, interesting information is given by an article [8]. Bernoulli devoted a great amount of time to the problem of spirals and wrote a treatise called *Spiral mirabilis (The marvelous spiral*). He studied the spiral many hours and finally understood that this spiral's qualities are nearly magic. Therefore, he wanted such a spiral engraved on his headstone. Unfortunately, by error, an Archimedean spiral was placed there instead of the logarithmical one, which he loved so much.

Another interesting fact dealing with the logarithmic spiral is its connection with the golden ratio. The definition of the golden ratio and its properties can be found e.g. in [13]. In the article [6] there is stated the following: A golden rectangle is a rectangle where the ratio of its length to its width equals to the golden ratio. If we cut away a square from this rectangle, we will get a rectangle, which is also golden. If we continue in cutting away in the same way, we will always get a golden rectangle. If we draw into any pair of the mother rectangle and the daughter rectangle two diagonals according to the following picture (See [6]), all these diagonals will intersect in one point. The series of the decreasing rectangles converge to this point.

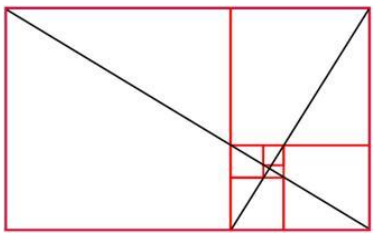


Fig. 3: Golden rectangles and their diagonals
Source: [6]

If we now draw a curve joining the points which divide the longer sides of the decreasing golden rectangles in the golden ratio (See Fig. 4), we will get a logarithmic spiral. The above-mentioned point is the pole of the spiral.

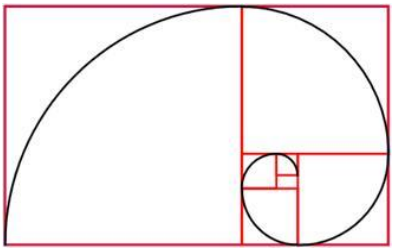


Fig. 4: Logarithmic spiral and a golden ratio
Source: [6]

Now we will deal with the logarithmic spiral from the mathematical point of view (according to [1] and [5]).

In the polar system of coordinates the logarithmic spiral can be expressed by the equation $r = a\, e^{b\varphi}$. The equation of the logarithmic spiral could be also written parametrically. For the coordinates of any point $X = [x, y]$ of a logarithmic spiral there applies (for this purpose we will denote the angle φ more conveniently as t)

$$X = r \cos t = a\, e^{bt} \cos t,\ y = r \sin t = a\, e^{bt} \sin t.$$

Let us note that for values b nearing zero, the spiral will near a circle. Let us admit the case $b = 0$, then the spiral turns into a circle. The change of the length of the radius vector can be expressed by a derivation

$$\frac{dr}{dt} = abe^{b\varphi} = br.$$

The growth of the spiral depends only on the value b, while the parameter a determines the distance of the origin of the spiral from its pole. Let us work out the length s of the logarithmic spiral from its origin $P = [a, 0]$ to any of its point $Q = [r, \alpha]$. The desired length s will be calculated with the use of the following integral (the details of the computation are omitted):

$$\int_0^{\alpha} \sqrt{r^2 + \left(\frac{dr}{d\varphi}\right)^2}\, d\varphi = \int_0^{\alpha} ae^{b\varphi} \sqrt{1+b^2}\, d\varphi = \frac{a\sqrt{1+b^2}}{b} \left[e^{b\varphi}\right]_0^{\alpha} = \frac{a\sqrt{1+b^2}}{b} \left(e^{b\alpha} - 1\right) = \frac{\sqrt{1+b^2}}{b} (r - a).$$

Similarly, we can derive the formula for the length of the logarithmic spiral between two points $P = [r, \alpha]$ and $Q = [s, \beta]$ (without detriment to generality, let us assume $\beta > \alpha$, therefore $s > r$ too):

$$s = \frac{\sqrt{1+b^2}}{b} (s - r).$$

3. HYPERBOLIC SPIRAL

Hyperbolic spiral was discovered by Pierre Varignon in 1704. It was further studied by Johann Bernoulli between 1710 and 1713 and also by Roger Cotes in 1722 ([5]). Hyperbolic spiral is a plane curve for which the product of the radius vector r and argument φ (i.e. the product of polar coordinates) is constant (See [1]). It is inverse to Archimedean spiral, i.e. in polar coordinates it has the equation

$$r = \frac{a}{\varphi}\ ,\ a \in \boldsymbol{R},\ a > 0,\ \varphi > 0.$$

The asymptotic point (called also a pole) is in its basic position in the Cartesian coordinate system the origin $O = [0, 0]$. The spiral starts in the infinite distance from its pole, approaches it and winds around it in tighter loops. See the following picture (taken from [10]).

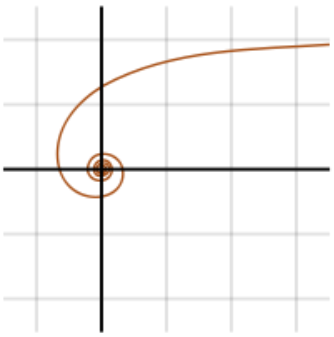


Fig. 5: Hyperbolic spiral
Source: [10]

The shape of the hyperbolic spiral can be shown more clearly if we rewrite the equation to the form $r\varphi = a$. Since the number a is a positive parameter, after its fixed selection we can write $r\varphi = const$. If the values of the angle φ approach zero, the points of the spiral approach a straight line parallel to a polar axis (axis x) whose distance from the polar axis is equal a. The straight line $y = a$ is the asymptote of the hyperbolic spiral $r = \frac{a}{\varphi}$. In the parametric formulation the hyperbolic spiral is defined by equations

$$x = \frac{a}{t}cost\ ,\ y = \frac{a}{t}sint$$

where $X = [x, y]$ is an arbitrary point of the hyperbolic spiral and the value of the angle φ is denoted as t. We will outline the calculation of the length of the hyperbolic spiral arc between its two points $P = [r, \alpha]$, $Q = [s, \beta]$ (without detriment to generality, let us assume that $\beta > \alpha$). At first let us calculate the derivation $\frac{dr}{d\varphi} = -\frac{a}{\varphi^2}$. Then there applies

$$s = \int_{\alpha}^{\beta} \sqrt{r^2 + \left(\frac{dr}{d\varphi}\right)^2}\, d\varphi = \int_{\alpha}^{\beta} \sqrt{\left(\frac{a}{\varphi}\right)^2 + \left(-\frac{a}{\varphi^2}\right)^2}\, d\varphi = a\int_{\alpha}^{\beta} \frac{1}{\varphi^2}\sqrt{1+\varphi^2}\, d\varphi.$$

The last integral is very difficult and leads to complicated relations. Therefore, we will not provide it here. Those interested could try to solve the indefinite integral $\int \frac{\sqrt{1+x^2}}{x^2}\, dx$ and thus see how complicated this problem is. We will present the theory necessary for solving the last indefinite integral which is a binomial integral (See e.g. an electronic source [14]).

Binomial integrals $\int x^m \left(a+bx^n\right)^p \, dx$, *(m, n, p are rational numbers), can be transformed to integrals of rational functions if at least one of numbers* $p, \frac{m+1}{n}, \frac{m+1}{n}+p$ *is an integer.*

a) If p is an integer, then according to the binomial theorem we will expand $\left(a+bx^n\right)^p$ *to a series, we will multiply individual members of the series by* x^m *and integrate; if the number* p *is a negative integer, we will find the least common multiple s of the denominators m* and *n and we will introduce the substitution* $x = t^s$.

b) If $\frac{m+1}{n}$ *is an integer, we will use the substitution* $a+bx^n = t^s$ *(s is the denominator of fraction p). Depending on the exponent of the integrated function after the substitution, i.e. if it is positive or negative, we will choose the further procedure, as is given in a).*

c) If $\frac{m+1}{n}+p$ *is an integer, we will rearrange the expression* $a+bx^n$ *by taking out* x^n *;we will obtain* $\left(a+bx^n\right) = x^n\left(ax^{-n}+b\right)$. *Then we will use the substitution* $ax^{-n}+b = t$.

If we use the above described theory for calculating the integral $\int \frac{\sqrt{1+x^2}}{x^2} \, dx$, we will see that with denoting $m = -2$, $n = 2$, $p = \frac{1}{2}$, $a = b = 1$, the solved integral is the integral of type c), i.e. the most complicated one of the three possible cases. The instruction for its calculation is given above, so we will not deal with it in this article further.

4. FERMAT SPIRAL

Fermat or parabolic spiral is first mentioned in 1636 in the writing of French mathematician Pierre de Fermata (1601–1665) *Ad locos planos et solidos lisagoge* (The introduction to the study of plane and solid curves), see [8]. This spiral is described by an equation

$$r^2 = a^2\varphi \, , \, a \in \mathbf{R}, \, a > 0, \, \varphi \geq 0.$$

For each positive value of the argument φ there exist two corresponding values of the radius vector r – the positive and the negative ones. Therefore, the resulting spiral is symmetric about the line $y = -x$ and it is depicted in the following picture (taken from [8]).

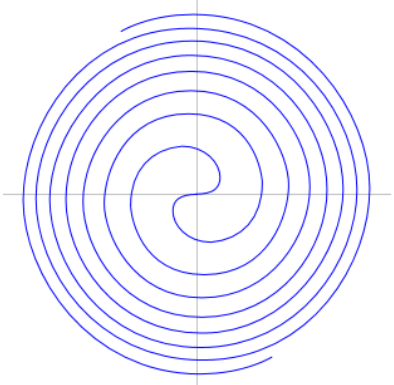


Fig. 6: Fermat spiral
Source: [8]

Fermat spiral is often described by equation $r = a\sqrt{\varphi}$; in this case only one half of the spiral is drawn. Sometimes it is expressed in the parametric form ($X = [x, y]$ is an arbitrary point of Fermat spiral and the value of the angle φ is denoted as t), i.e.

$$x = a\sqrt{t}\ \cos t, y = a\sqrt{t}\ \sin t\ , t > 0.$$

Both halves of Fermat spiral are drawn in different colours in the following picture (taken from [17]).

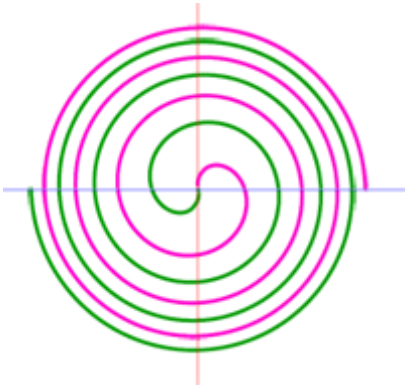


Fig. 7: Fermat spiral with colour differentiation
Source: [17]

The length of the Fermat spiral between its two points will not be given in this article. The reason is following: although we will use the definitional formula $r = a\sqrt{\varphi}$ for one half of Fermat spiral, after substitution to a general formula $s = \int_{\alpha}^{\beta} \sqrt{r^2 + \left(\frac{dr}{d\varphi}\right)^2}\, d\varphi$ and after its rearrangement, we will get an integral $s = \frac{a}{2}\int_{\alpha}^{\beta} \sqrt{\frac{4\varphi^2 + 1}{\varphi}}\, d\varphi$. This integral is binomial, but it is not one of the three above given cases, when the binomial integral can be transformed into an integral of some rational function. Its value can be determined numerically for the given limits, but it is not to be mentioned in this article.

Let us note that Fermat spiral is a special case of spirals of higher orders. Spirals of higher orders are spirals with polar equations

$$r^m = a^{-m} \cdot \frac{\varphi}{2\pi}, \text{ where } a, m \text{ are constants.}$$

Special cases of spirals of higher orders:

$r^2 = a^2\varphi$ Fermat spiral

$r^2 = a\varphi^2 - l$ Galileo spiral – it represents the trajectory of a mass point which falls freely with respect to the rotating Earth

$(r - a)^2 = 2ap\varphi$ parabolic spiral

Galileo spiral and parabolic spiral will not be dealt with in this article; details can be found e.g. in an electronic source [15].

5. LITUUOV SPIRAL

Lituuov spiral was proposed by Cotes in 1722 (lituus means a hook, [15]). Maclaurin used this expression in the book Harmonia Mensurarumin in 1722. It is an inverse spiral to Fermat spiral, i.e. it is a spiral with an equation $r = \frac{a}{\sqrt{\varphi}}$, $\varphi > 0$. The illustration is in the following picture (taken from [16]).

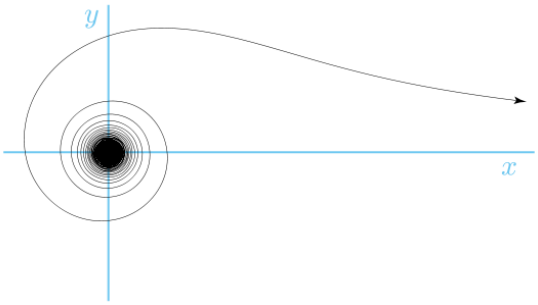


Fig. 8: Lituuov spiral ($r > 0$)
Source: [16]

In parametric formulation we will determine the coordinates of an arbitrary point of this spiral using relations $x = \frac{a}{\sqrt{t}} \cos t$, $y = \frac{a}{\sqrt{t}} \sin t$, $t > 0$.

Sometimes Lituuov spiral is expressed in the shape $r^2 = \frac{a}{\varphi}$. In this case the spiral has two branches depending on the sign of number r. In Figure 8 there is the branch for positive r. Both branches are depicted in Figure 9 (taken from [17]).

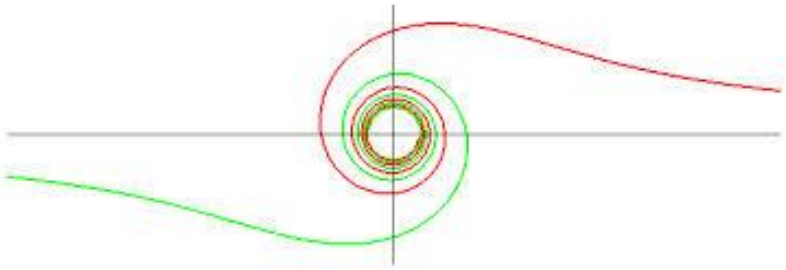


Fig. 9: Lituuov spiral – both branches
Source: [17]

Because of the difficulty while calculating the finite integral, we will not deal with the length of Lituuov spiral between its two points as well.

CONCLUSION

If we write the equation of the spiral in the more general form $r = a\,\varphi^b$, then we will get ([7]): for $b = 1$ Archimedean spiral with the equation $r = a\,\varphi$,

for $b = -1$ hyperbolic spiral with the equation $r = \frac{a}{\varphi}$, (an inverse curve to Archimedean spiral),

for $b = \frac{1}{2}$ Fermat spiral with the equation $r = a\sqrt{\varphi}$,

for $b = -\frac{1}{2}$ Lituuov spiral with the equation $r = \frac{a}{\sqrt{\varphi}}$, (an inverse curve to Fermat spiral).

Another interesting mathematics area is a theory of sinusoidal spirals (although they are not actually spirals). These spirals were at first dealt with by Colin Maclaurin. The sinusoidal spiral can be described by one of equivalent equations (it depends on the rotation of the coordination system)

$$r^m = a^m \sin(m\varphi),\ r^m = a^m \cos(m\varphi),\ \text{kde } m \in \mathbf{Q},\ \varphi > 0.$$

The theory and description of sinusoidal spirals can be the topic of further articles. Therefore let us only mention the most often occurring cases.

For $m = 1$ it is a circle, for $m = 2$ it is a lemniscate of Bernoulli,

for $m = -1$ we get a line, for $m = -2$ it is a rectangular hyperbola,

for $m = \frac{1}{2}$ it is a cardioid and for $m = -\frac{1}{2}$ it is a parabola.

A detailed information about these geometric figures and other technical curves as well can be found in a synoptic publication [1].

The theory of spirals and technical curves forms the part of Mathematics curriculum at technical universities where there are a number of theoretical books on this topic (e.g. [3]). At other universities, especially while teaching future Mathematics teachers, this theory is not dealt with and there is neither suitable nor accessible literature for them. The response to this situation is this article whose aim is to supply some examples of accessible literature for them. In conclusion, let us again mention that this whole part of Mathematics (with connection to the history of Mathematics, mathematical analysis and geometry) can represent a significant motivational aspect ([2]). What is more, this topic is not too distant from the secondary school Mathematics although it contains some parts from the higher stage. Therefore, while students struggle to penetrate this topic, they can devote their energies to studying Mathematics texts, especially mathematical analysis.

REFERENCES:

[1] BARTSCH, H., J. *Matematické vzorce.* Praha : SNTL, 1987. 832 pp.

[2] BERÁNEK, Jaroslav *Spirály v matematice*. In NOVOTNÁ, J. *Motivace nadaných žáků a studentů v matematice a přírodních vědách*. Brno : Masarykova Univerzita, 2015. pp. 5–15. ISBN 978-80-210-8146-8.

[3] BITTNEROVÁ, D. *Applications od GeoGebra for calculations figure areas bounded by roses*. In *MITAV 2019, post-conference proceedings of extended versions of selected papers*. Brno : University of Defence in Brno, 2019. ISBN 978-80-7582-123-2

[4] BUDINSKÝ, B., KEPR, B. *Základy diferenciální geometrie s technickými aplikacemi*. Praha : SNTL, 1970.

[5] JAREŠOVÁ, M., VOLF, I. *Matematika křivek.* Study text for Physical Olympics participants and others interested in physics. Hradec Králové : Nakladatelství MAFY, 2009. 64 pp. Available from http://fyzikalniolympiada.cz/texty/matematika/mkrivek.pdf (retrieved on 6. 4. 2021).

[6] REICHL, J. *Zlatý obdélník a logaritmická spirála*. Available from http://fyzika.jreichl.com/main.article/view/1465-zlaty-obdelnik-a-logaritmicka-spirala (retrieved on 7. 4. 2021).

[7] ŠKRÁŠEK, J., TICHÝ, Z. *Základy aplikované matematiky.* Praha : SNTL, 1986.

[8] *Spirály.* Available from http://www.eprojekt.gjs.cz/Services/Downloader.ashx?id=13159, (retrieved on 6. 4. 2021).

[9] *Archimédova spirála*. Available from https://cs.wikipedia.org/wiki/Archim%C3%A9dova_spir%C3%A1la (retrieved on 4. 4. 2021).

[10] *Hyperbolická spirála.* Available from https://cs.wikipedia.org/wiki/Hyperbolick%C3%A1_spir%C3%A1la (retrieved on 3. 4. 2021).

[11] *Logaritmická spirála.* Available from http://cs.wikipedia.org/wiki/Logaritmicka_spirala (retrieved on 7. 4. 2021).

[12] *Archimédova spirála.* In *Ottův slovník naučný*. Available from http://leccos.com/index.php/clanky/archimedova-spirala (retrieved on 5. 4. 2021).

[13] *Zlatý řez.* Available from https://cs.wikipedia.org/wiki/Zlat%C3%BD_%C5%99ez (retrieved on 31. 3. 2021).

[14] *Binomial Integral.* Available from http://www.nabla.hr/CL-IndefIntegralB5.htm (retrieved on 2. 4. 2021).

[15] *Spirály.* Available from http://www.matematika.cz/content/rovinne-krivky/spiraly/spiraly.doc (retrieved on 8. 4. 2021)

[16] *Lituus (mathematics).* Available from http://en.wikipedia.org/wiki/Lituus_(mathematics) (retrieved on 8. 4. 2021).

[17] *Lituus spiral equation*. Available from http://xahlee.info/SpecialPlaneCurves_dir/Lituus_dir/lituus.html (retrieved on 3. 4. 2021).

VOLUMES OF SOME SOLIDS OF REVOLUTION AND APPLICATIONS IN GEOGEBRA

Daniela Bittnerová
Technical University of Liberec
Studentská 2, 461 17 Liberec, daniela.bittnerova@tul.cz

Abstract: *The paper presents an alternative technic of calculation volumes of solids of revolution, and also a possibility of using the special geometric software GeoGebra 5.0 for some applications of them.*

Keywords: solid of revolutions, parametric equations, GeoGebra

INTRODUCTION

The contribution follows on from the papers published by the author in the conferences AMEE´13 and AMEE´16 where volumes of solids as a topological problems were discussed (see [1] , [2], and [3]). An alternative method of calculating areas and/or volumes was proved generally in $\mathbf{E_n}$ there. The method leads to an easier calculation because integrals of *n*-1 dimension is used instead of the *n*-dimensional (in $\mathbf{E_3}$ - double integrals instead of the triple ones), which is commonly found in conventional examples.

Now we are interesting in special cases of solids of revolution, and also we present a view of a possibility, how to use the dynamic geometric software of GeoGebra to calculations and demonstrations of them. The case of some closed curves has been published in MITAV 2019 – in [4]. Geo-Gebra is very popular also between teachers and students because it provides many interesting options for solving examples and demonstrations tasks. The big advantage is also that this software is free.

1 CALCULATION VOLUMES USING AN ALTERNATIVE METHOD IN $\mathbf{E_3}$

An alternative method of calculating the volume is based of knowledge of the parametric description of surface of the body. The problem of parametric descriptions of the surface areas of solid is investigated as a problem in the topological sense. There was also proved the formula for the calculation of the volume in *n*-dimensional space for the case that the surface areas are the smooth, resp. piecewise smooth areas in the Euclidean space of the corresponding dimensions.

Let $X = [x_1, x_2, x_3\ ]$ be point and its Cartensian coordinates in $\mathbf{E_3}$, $U = [u_1, u_2\ ]$ the Cartensian coordinates of the point in $\mathbf{E_2}$, $\boldsymbol{\Omega}$ the bounded closed domain in $\mathbf{E_2}$, $x_i(u_1, u_2), i = 1, 2, 3,$ given functions defined on some domain $\mathbf{O} \subset \mathbf{E_2}$, $\boldsymbol{\Omega} \subset \mathbf{O}$.

Let us suppose that

- The vector function $x(u)$ has almost everywhere in $\boldsymbol{\Omega}$ the continuous partial derivatives

$$\frac{\partial x_i}{\partial u_j} \quad \text{for } i = 1, 2, 3, j = 1, 2;$$

- The rank of the matrix $\left(\frac{\partial x_i}{\partial u_j}\right)_{(3\times 2)}$ is equal to 2 almost everywhere in $\mathbf{\Omega}$;
- The subset $\mathbf{P^0} = \{x \in \mathbf{E_3}; x = x(u), u \in \Omega\}$ of the set $\mathbf{P} = \{x \in \mathbf{E_3}; x = x(u), u \in \mathbf{\Omega}\}$ is a homeomorphic range of the set int $\mathbf{\Omega}$ in $\mathbf{E_3}$.

Then, the closure *W* of the set **P** is the boundary of the 3-dimensional solid in the space $\mathbf{E_3}$. The volume $V = \mu W$ of the 3-dimensional solid can be calculated by the formula (see [1] and [2]).

$$V = \frac{1}{3}\iint_{\Omega} \Delta(u)\, \mathrm{d}u_1\, \mathrm{d}u_2 \tag{1}$$

where

$$\Delta(u) = \begin{vmatrix} x_1(u) & x_2(u) & x_3(u) \\ \frac{\partial x_1}{\partial u_1} & \frac{\partial x_2}{\partial u_1} & \frac{\partial x_3}{\partial u_1} \\ \frac{\partial x_1}{\partial u_2} & \frac{\partial x_2}{\partial u_2} & \frac{\partial x_3}{\partial u_2} \end{vmatrix}. \tag{2}$$

2 CLASSICAL AND ELLIPTICAL TOROID

A toroid is a surface of revolution generated by revolving a circle in $\mathbf{E_3}$, about an axis that is coplanar with the circle and has no common point with it and has no common point with it (Fig. 1 and 2).

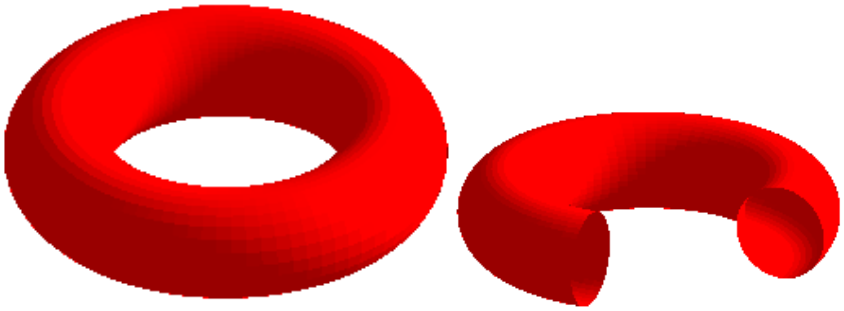


Fig. 1 and Fig. 2 A toroid and a part of a toroid

Parametric equations of the surface are

$$\begin{aligned} x_1 &= (a + r\cos v)\cos u \\ x_2 &= (a + r\cos v)\sin u \\ x_3 &= r\sin v \\ v &\in \langle 0; 2\pi\rangle, u \in \langle 0; 2\pi\rangle \end{aligned} \tag{3}$$

where r denotes the radius of the rotating circle $k\,(S; r)$ and $a > r$ is a distance of the centre $S = [a; 0; 0]$ from the z-axis of rotation.
Then according to the symmetry, the formula (2) is of the form

$$\Delta(u,v) = \begin{vmatrix} (a + r\cos v)\cos u & (a + r\cos v)\sin u & r\sin v \\ -(a + r\cos v)\sin u & (a + r\cos v)\cos u & 0 \\ -r\sin v\ \cos u & -r\sin v\ \sin u & r\cos v \end{vmatrix} = \tag{4}$$

$$= \lceil \text{a calculation of the determinant} \rfloor =$$

$$= r^2\sin^2 v(a + r\cos v)(\sin^2 u + \cos^2 u) + r\cos v\,(a + r\cos v)^2(\cos^2 u + \sin^2 u) =$$

$$= r(a + r\cos v)[r\sin^2 u + \cos v(a + r\cos v)] =$$

$$= r(a + r\cos v)[r\sin^2 u + a\cos v + r\cos^2 v] =$$

$$= r(a + r\cos v)(r + a\cos v) = r(ar + r^2\cos v + a^2\cos v + ar\cos^2 v) =$$

$$= r\left(ar + (r^2 + a^2)\cos v + ar\cos^2 v\right)$$

and the volume is equal to

$$V = \tfrac{4}{3}r\int_0^{\pi} \mathrm{d}u \int_0^{\pi}\left(ar + (r^2 + a^2)\cos v\ \cos v + ar\cos^2 v\right)\mathrm{d}v = \tag{5}$$

$$= \frac{4}{3}r\pi\left[arv + (r^2 + a^2)\sin v + ar\cdot\frac{v + \sin v\ \cos v}{2}\right]_0^{\pi} =$$

$$= \frac{4}{3}r\pi\left(ar\pi + \frac{ar\pi}{2}\right) = \frac{4}{3}r^2\pi^2\frac{3}{2}a = 2ar^2\pi^2$$

In the case of **solids of revolution**, we can simplify the formulas (3 - 5). Using the symmetry of the toroid, we can calculate the volume of only that part which lies above the plane (xy). Parametric equations of the semi-surface are:

$$\begin{aligned} x_1 &= f(u)\cos v = u\cos v \\ x_2 &= f(u)\sin v = u\ \sin v \\ x_3 &= g(u) = \sqrt{r^2 - (u - a)^2} \\ v &\in \langle 0; 2\pi\rangle, u \in \langle a - r; a + r\rangle \end{aligned} \tag{6}$$

The determinant:

$$\Delta(u,v) = \begin{vmatrix} f(u)\cos v & f(u)\sin v & g(u) \\ f'(u)\cos v & f'(u)\sin v & g'(u) \\ -f(u)\sin v & f(u)\cos v & 0 \end{vmatrix} = \tag{7}$$

$$= g(u)\left[f'(u)f(u)\cos^2 v + f(u)f'(u)\sin^2 v\right] - g'(u)[f^2(u)\cos^2 v + f^2(u)\sin^2 v] =$$

$$= g(u)f'(u)f(u) - g'(u)f^2(u) = f(u)f'(u)g(u) - g'(u)f^2(u)$$

In this particular case, we have got

$$\Delta(u) = u\sqrt{r^2 - (u-a)^2} + \frac{(u-a)u^2}{\sqrt{r^2-(u-a)^2}} \tag{8}$$

The volume:

$$V = 2\frac{2\pi}{3}\int_{r-a}^{r+a}\Delta(u)\,\mathrm{d}u = \tag{9}$$

$$= \frac{4\pi}{3}\left[\frac{3ar^2}{2}\arcsin\frac{u-a}{r} - \left(\frac{a^2}{2} + r^2 - \frac{au}{2}\right)\sqrt{r^2-(u-a)^2}\right]_{a=r}^{a+r} = \cdots = 2ar^2\pi^2$$

3 VOLUME OF ELLIPTICAL RINGS

An elliptical ring with a circular cross-section is a surface formed by moving a circle *k* (*S*,*r*) in the direction of an elliptical orbit that is perpendicular to the circle *k* (Fig. 3).

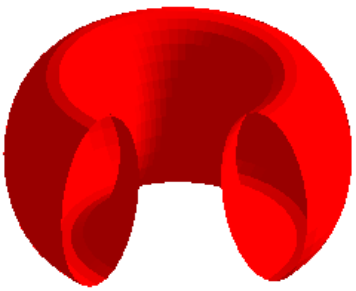


Fig. 3 A part of an elliptical ring with a circular cross-section

As a circle $k(S,r)$ can be taken as a special case of an ellipse if $c = d = r$, we can use the previous results, i.e.:

Parametric equations of the surface:

$$\begin{aligned} x_1 &= (a + r\cos v)\cos u \\ x_2 &= (b + r\cos v)\sin u \\ x_3 &= r\sin v \end{aligned} \tag{10}$$

$$v \in \langle 0; 2\pi\rangle, u \in \langle 0; 2\pi\rangle$$

The determinant:

$$\Delta(u,v) = \begin{vmatrix} (a + r\cos v)\cos u & (b + r\cos v)\sin u & r\sin v \\ -(a + r\cos v)\sin u & (b + r\cos v)\cos u & 0 \\ -r\sin v\ \cos u & -r\sin v\ \sin u & r\cos v \end{vmatrix} = \cdots = \quad (11)$$

$$= abr\cos v + r^2(a+b)\cos^2 v + ar^2\sin^2 u\sin^2 v + r^3\sin^2 v\cos v + br^2\sin^2 v\cos^2 u.$$

and the volume is equal to

$$V = \tfrac{1}{3}\int_0^{2\pi}\int_0^{2\pi}\Delta(u)\,\mathrm{d}u\,\mathrm{d}v = (a+b)\,r^2\pi^2. \quad (12)$$

An elliptical ring with an elliptical cross-section is a surface formed by moving an ellipse *e* (*S*, *c*, *d*) in the direction of an elliptical orbit that is perpendicular to the ellipse *e*.

Parametric equations of the surface:

$$\begin{aligned} x_1 &= (a + c\cos v)\cos u \\ x_2 &= (b + c\cos v)\sin u \\ x_3 &= d\sin v \\ v &\in \langle 0; 2\pi\rangle, u \in \langle 0; 2\pi\rangle \end{aligned} \quad (13)$$

where *a* and *b* denote the semi-axes of the trajectory of the moving ellipse *e* (*S*, *c*, *d*). We suppose that the surface is not intersecting, i. e. min (*a*, *b*) ≥ *c*.
In particular case, the formula (2) takes the form

$$\Delta(u,v) = \begin{vmatrix} (a + c\cos v)\cos u & (b + c\cos v)\sin u & d\sin v \\ -(a + c\cos v)\sin u & (b + c\cos v)\cos u & 0 \\ -c\sin v\ \cos u & -c\sin v\ \sin u & d\cos v \end{vmatrix} = \cdots = \quad (14)$$

$$= abc\cos v + cd(a+b)\cos^2 v + acd\sin^2 u\sin^2 v + c^2 d\sin^2 v\cos v + bcd\sin^2 v\cos^2 u.$$

The volume of ring is equal to

$$V = \tfrac{1}{3}\int_0^{2\pi}\int_0^{2\pi}\Delta(u)\,\mathrm{d}u\,\mathrm{d}v = (a+b)\,cd\pi^2. \quad (15)$$

4 VOLUME OF AN AXOID (HORN TOROID)

An axoid (a horn toroid) is a surface formed by revolving a circle k (S,r) in $\mathbf{E}_3$ about an axis that is a tangent to the circle. In this case, the circle rotates about the z-axis. (Fig. 4):

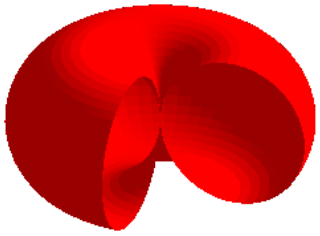


Fig. 4 The part of the axoid

The general alternative method:

Parametric equations of the surface:

$$\begin{aligned} x_1 &= r(1+\cos v)\cos u \\ x_2 &= r(1+\cos v)\sin u \\ x_3 &= r\sin v \\ v &\in \langle 0; 2\pi\rangle, u \in \langle 0; 2\pi\rangle \end{aligned} \tag{16}$$

The determinant:

$$\Delta(u,v) = \begin{vmatrix} r(1+\cos v)\cos u & r(1+\cos v)\sin u & r\sin v \\ -r(1+\cos v)\sin u & r(1+\cos v)\cos u & 0 \\ -r\sin v\ \cos u & -r\sin v\ \sin u & r\cos v \end{vmatrix} = \tag{17}$$

$$= \cdots = r^3(1+\cos v)^2$$

The volume is equal to

$$V = \tfrac{4}{3}r^3 \int_0^{\pi} \mathrm{d}u \int_0^{\pi} [r^3(1+\cos v)^2]\,\mathrm{d}v = \cdots = 2r^3\pi^2. \tag{18}$$

The special method for solids of revolution:

The parametric equations:

$$\begin{aligned} x_1 &= f(u)\cos v = u\cos v \\ x_2 &= f(u)\sin v = u\ \sin v \\ x_3 &= g(u) = \sqrt{r^2-(u-r)^2} \end{aligned} \tag{19}$$

$$v \in \langle 0; 2\pi\rangle, u \in \langle 0; 2r\rangle$$

The determinant is equal to:

$$\Delta(u) = u\sqrt{r^2 - (u-r)^2} + \frac{(u-r)u^2}{\sqrt{r^2-(u-r)^2}} \tag{20}$$

The volume:

$$V = 2\frac{2\pi}{3}\int_0^{2r} \Delta(u)\,\mathrm{d}u = \tag{21}$$

$$= \frac{4\pi}{3}\left[\frac{3r^3}{2}\arcsin\frac{u-r}{r} - \left(\frac{3r^2}{2} - \frac{ru}{2}\right)\sqrt{r^2-(u-r)^2}\right]_0^{2r} = \cdots = 2r^3\pi^2$$

A melanoid (a spindle toroid) is a surface in **E**$_3$ formed by revolving a circle k (S,r) about an axis that is a chord to the circle. In this case, the circle rotates about the z-axis. (Fig. 5).

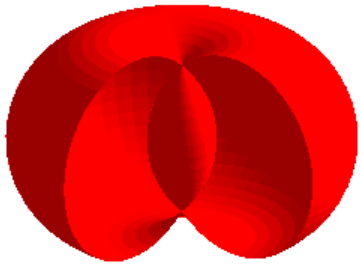


Fig. 5 The part of the melanoid

The special method for solids of revolution:

We can use the same parametric equations as for the axoid (19). In this case, the calculation of the volume must be divided to two steps:

The volume of the exterior part (an apple surface):

$$V_o = 2\frac{2\pi}{3}\int_0^{a+r} \Delta(u)\,\mathrm{d}u = \tag{22}$$

$$= \frac{4\pi}{3}\left[\frac{3ar^2}{2}\arcsin\frac{u-a}{r} - \left(\frac{a^2}{2} + r^2 - \frac{au}{2}\right)\sqrt{r^2-(u-a)^2}\right]_0^{a+r} =$$

$$= \frac{4\pi}{3}\left[\frac{3ar^2}{2}\left(\frac{\pi}{2} + \arcsin\frac{a}{r}\right) - \left(\frac{a^2}{2} + r^2\right)\sqrt{r^2 - a^2}\right]$$

The volume V_i of the interior part (a lemon surface):

$$V_i = 2\frac{2\pi}{3}\int_{a-r}^{0} \Delta(u)\,\mathrm{d}u = \tag{23}$$

$$= \frac{4\pi}{3}\left[\frac{3ar^2}{2}\arcsin\frac{u-a}{r} - \left(\frac{a^2}{2} + r^2 - \frac{au}{2}\right)\sqrt{r^2 - (u-a)^2}\right]_0^{a+r} =$$

$$= \frac{4\pi}{3}\left[\frac{3ar^2}{2}\left(\frac{\pi}{2} - \arcsin\frac{a}{r}\right) + \left(\frac{a^2}{2} + r^2\right)\sqrt{r^2 - a^2}\right]$$

The volume *V* of the "hollow" body is determined by the difference between these two volumes, i.e.:

$$V = V_o - V_i = 4ar^2\pi \arcsin\frac{a}{r} \tag{24}$$

By comparing both method, we can see, that for each special case, some method is more suitable.

5 USING GEOGEBRA FOR COMPUTING VOLUMES FOR A TOROID

The volumes of solids of revolution are very easy to count using the software known as GeoGebra. For example, the volume of the toroid is computed by the formula above – see the applet in the Fig. 6. There are used the sliders for setting the values of the parameters *R* (the distance from the centre of the rule of the tube to he centre of the toroid, and *r* (the radius of the tube) of the toroid in the applet. Changing the values of the parameters, we create the toroid of the different parameters. It is possible to construct the analogous applets for computing the volumes of the axoid, respective melanoid.

CONCLUSION

In the real practice, we are able to find many technical applications where the presented method could be used. The software GeoGebra is a very good tool to the demonstrations of it. Its advantage is that it is available for free and that it is used a lot of when teaching in schools. Intermediate calculations of determinants and volumes can be easy to obtain.
All figures were made by the authors. Some of them have been published in [2].

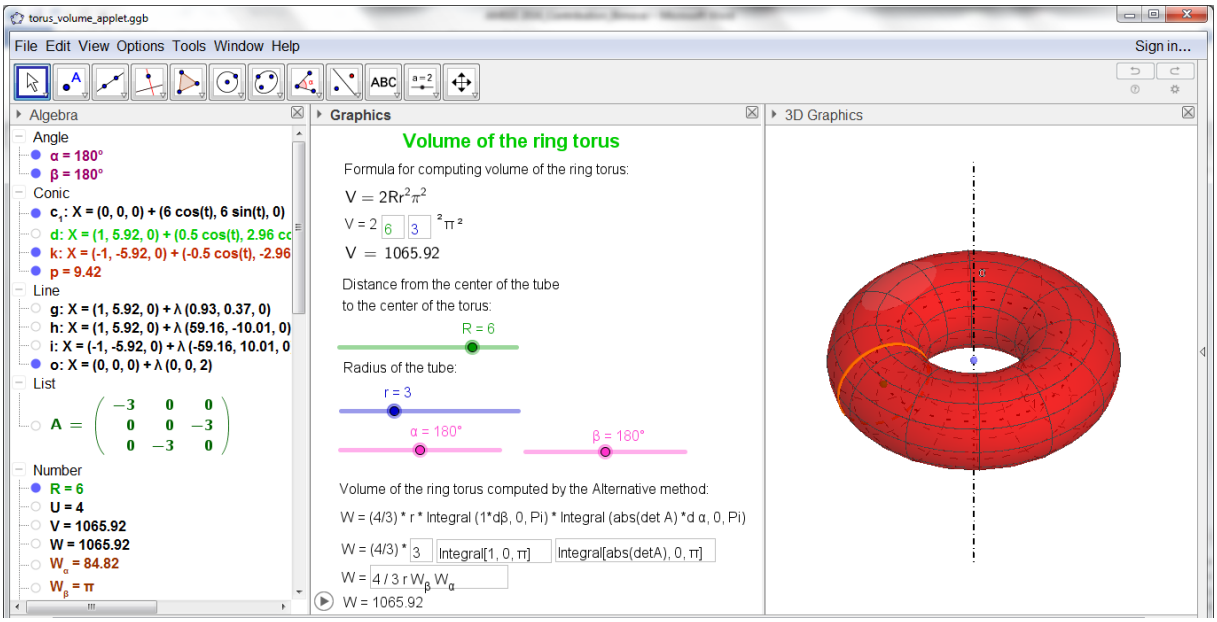


Fig. 6 The applet for computing volume of a toroid created in GeoGebra 5.0 [2]

References

[1] Bittnerova, D. Alternative Method for Calculations of Volumes by Using Parameterizations Surface Areas. In: *Applications of Mathematics in Engineering and Economics, AIP Conference Proceedings 1570, 2013*. Sozopol: Technical University Sofia 2013, pp. 3-10. DOI: 10.1063/1.4854736

[2] Bittnerova, D., Bímová, D. Volume of Torus and His Applications Unconventionally In: *Applications of Mathematics in Engineering and Economics*, AIP Conference Proceedings 1789, 2016. Sozopol: Technical University Sofia 2016. DOI: 10.1063/1.4968486

[3] Bittnerova, D., Bímová, D. Some Applications of Unconventional Methods for Volumes. In: *Applications of Mathematics in Engineering and Economics, AIP Conference Proceedings 1789, 2016*, Sozopol: Technical University Sofia 2016. DOI: 10.1063/1.4968486

[4] Bittnerova, D. Applications of GeoGebra for Calculations Figure Areas Bounded by Roses. In: *Mathematics, Information Technologies and Applied Sciences 2019, post-conference proceedings of extended versions of selected papers*. Brno: University of Defence, 2019, pp. 30 – 37. [Online]. Available at: <http://mitav.unob.cz/data/MITAV 2019 Proceedings.pdf>. ISBN 978-80-7582-123-2

ON A DISCRETE VARIANT OF THE EMDEN-FOWLER EQUATION

Josef Diblík
Brno University of Technology, Faculty of Civil Engineering,
Department of Mathematics and Descriptive Geometry,
diblik.j@fce.vutbr.cz
Faculty of Electrical Engineering and Communication, Department of Mathematics,
diblik@feec.vutbr.cz
Brno, Czech Republic

Evgeniya Korobko
Brno University of Technology, Faculty of Electrical Engineering and Communication,
Department of Mathematics,
xkorob01@stud.feec.vutbr.cz, jakovi300195@yandex.ru
Brno, Czech Republic

Abstract: *In the paper, the discrete Emden-Fowler type equation*

$$\Delta^2 u(k) \pm k^\alpha u^m(k) = 0$$

is considered where $k \geq k_0$, k is an independent variable, k_0 is a fixed integer, $u\colon \{k_0, k_0+1, ...\} \to \mathbb{R}$, $\Delta u(k)$ is the first difference of $u(k)$, $\Delta^2 u(k)$ is the second difference of $u(k)$, m and α are real numbers. A result on asymptotic behaviour of solutions when $k \to \infty$ is proved and admissible values m and α satisfying assumptions of this result are considered in an (m, α)-plane.

Keywords: Emden-Fowler equation, discrete equation, nonlinear equation, asymptotic behaviour.

INTRODUCTION

Consider a discrete Emden-Fowler equation

$$\Delta^2 u(k) \pm k^\alpha u^m(k) = 0, \tag{1}$$

where $u\colon \mathbb{N}(k_0) := \{k_0, k_0+1, ...\} \to \mathbb{R}$, $\Delta u(k) = u(k+1) - u(k)$ is the first difference of $u(k)$, $\Delta^2 u(k) = u(k+2) - 2u(k+1) + u(k)$ is the second difference of $u(k)$ and m and α are real numbers. Throughout the paper we assume $\alpha \neq 0$ and $m \neq 0, 1$. A function $u\colon \mathbb{N}(k_0) := \{k_0, k_0+1, ...\} \to \mathbb{R}$ is called a solution of equation (1) if, for every $k \in \mathbb{N}(k_0)$, equation (1) is satisfied.

Equation (1) is a discrete variant of the well-known second-order differential Emden-Fowler equation [1]. In the previous research of the second author [5, Corollary 1], it is proved that if m and α satisfy either

$$0 < m < 1, \quad \alpha < -2 \tag{2}$$

or

$$m > 1, \quad -2 < \alpha < \frac{1}{2}\left(-(m-1) + \sqrt{(m-1)^2 + 16m}\right) \tag{3}$$

then the equation (1) has a solution with the asymptotic behavior

$$u(k) = \frac{a}{k^s} + \frac{b}{k^{s+1}} + O\left(\frac{1}{k^{s+\gamma+1}}\right) \tag{4}$$

when $k \to \infty$, where O is the Landay order symbol "big" O, $\gamma \in (0,1)$ is a fixed number and

$$s = \frac{\alpha+2}{m-1}\,, \quad a = [\mp s(s+1)]^{1/(m-1)}, \quad b = \frac{as(s+2)}{s+2-ms}\,. \tag{5}$$

In this article, we are going to supplement the set of previous conditions (2), (3) for the existence of the solution of the difference equation (1) having asymptotic behaviour (4) with an additional set of sufficient conditions such that asymptotic behaviour (4) will be preserved.

Remark 1. *Equation* (1) *splits into two equations*

$$\Delta^2 u(k) + k^\alpha u^m(k) = 0$$

and

$$\Delta^2 u(k) - k^\alpha u^m(k) = 0.$$

Nevertheless, above and in the the remaining part of the paper we apply the following restriction to the first of them, i.e., when in equation (1) *sign "+" is considered. The sign "+" in equation* (1) *is applicable only in the case of* m *having the form of a rational number,* $m = p/q$ *where* p *and* q *are integers, such that the difference* $p - q$ *is odd. Then equation* (1) *has the solution with asymptotic behaviour* (4)*.*

1 PRELIMINARIES

To prove the main result formulated below we need the following auxiliary result (we refer to original sources [2, 3]). Let a system of discrete equations

$$\Delta Y(k) = F(k, Y(k)), \quad k \in \mathbb{N}(k_0) \tag{6}$$

be given, where $Y = (Y_0, \ldots, Y_{n-1})^T$ and $F = (F_1, \ldots, F_n)^T \colon \mathbb{N}(k_0) \times \mathbb{R}^n \to \mathbb{R}^n$. A solution $Y = Y(k)$ of system (6) we define as a function $Y \colon \mathbb{N}(k_0) \to \mathbb{R}^n$ such that for each $k \in \mathbb{N}(k_0)$ equation (6) is satisfied. Let a point

$$Y(k_0) = Y^0, \quad Y^0 \in \mathbb{R}^n \tag{7}$$

be fixed. It is well-known that the initial problem (6), (7) determines a unique solution to (6). Let functions $b_i, c_i \colon \mathbb{N}(k_0) \to \mathbb{R}$, $i = 1, ..., n$ be fixed and satisfy

$$b_i(k) < c_i(k), \quad k \in \mathbb{N}(k_0), \quad i = 1, ..., n. \tag{8}$$

To formulate the auxiliary result, define functions $B_i, C_i \colon \mathbb{N}(k_0) \times \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, n$

$$B_i(k, Y) := -Y_{i-1} + b_i(k), \quad C_i(k, Y) := Y_{i-1} - c_i(k)$$

and sets

$$\Omega_B^i := \{(k,Y)\colon k \in \mathbb{N}(k_0), B_i(k,Y) = 0, B_j(k,Y) \le 0, C_p(k,Y) \le 0, \\ \forall\ j,p = 1,\dots,n,\ j \ne i\},$$

$$\Omega_C^i := \{(k,Y)\colon k \in \mathbb{N}(k_0), C_i(k,Y) = 0, B_j(k,Y) \le 0, C_p(k,Y) \le 0, \\ \forall\ j,p = 1,\dots,n,\ p \ne i\}$$

where $i = 1, ..., n$.

Lemma 1. *Let a function $F(k,Y)$ be continuous with respect to Y. If*

$$F_i(k,Y) < b_i(k+1) - b_i(k) \tag{9}$$

for every $i = 1,\dots,n$ and every $(k,Y) \in \Omega_B^i$ and

$$F_i(k,Y) > c_i(k+1) - c_i(k) \tag{10}$$

for every $i = 1,\dots,n$ and every $(k,Y) \in \Omega_C^i$, then there exists a solution $Y = Y(k)$, $k \in \mathbb{N}(k_0)$ of system (6) *such that*

$$b_i(k) < Y_{i-1}(k) < c_i(k), \quad k \in \mathbb{N}(k_0), \quad i = 1,\dots,n. \tag{11}$$

2 MAIN RESULT

In this part, we prove the main result of the contribution on the existence of a solution to equation (1) when $k \to \infty$. It is formulated in terms of coefficients s, m and some auxiliary constants γ and ε_i, $i = 1,2,3,4$.

Theorem 1. *Let either*

$$s > 0, \quad m > 0 \tag{12}$$

or

$$-1 < s < 0, \quad m < 0. \tag{13}$$

Assume that there exists a constant γ, satisfying $0 < \gamma < 1$ and positive numbers ε_i, $i = 1,2,3,4$, such that

$$\varepsilon_3 < \varepsilon_1 \cdot \frac{\gamma + s + 1}{s+1}, \tag{14}$$

$$\varepsilon_4 < \varepsilon_2 \cdot \frac{\gamma + s + 1}{s+1}, \tag{15}$$

$$\varepsilon_1 < \varepsilon_3 \cdot \frac{\gamma + s + 2}{ms}, \tag{16}$$

$$\varepsilon_2 < \varepsilon_4 \cdot \frac{\gamma + s + 2}{ms}. \tag{17}$$

Then, for sufficiently large fixed $k_0 > 0$, there exists a solution $u\colon \mathbb{N}(k_0) \to \mathbb{R}$ of equation (1) *such that, for every $k \in \mathbb{N}(k_0)$ asymptotic representation* (4) *holds or, more presisely, this solution satisfies*

$$-\frac{\varepsilon_1}{k^\gamma} < \left[u(k) - \frac{a}{k^s} - \frac{b}{k^{s+1}}\right]\left[\frac{b}{k^{s+1}}\right]^{-1} < \frac{\varepsilon_2}{k^\gamma}\,, \tag{18}$$

$$-\frac{\varepsilon_3}{k^\gamma} < \left[\Delta u(k) - \Delta\left(\frac{a}{k^s}\right) - \Delta\left(\frac{b}{k^{s+1}}\right)\right]\left[\Delta\left(\frac{b}{k^{s+1}}\right)\right]^{-1} < \frac{\varepsilon_4}{k^\gamma}\,, \tag{19}$$

$$-\frac{\varepsilon_1}{k^\gamma} + O\left(\frac{1}{k}\right) < \left[\Delta^2 u(k) - \Delta^2\left(\frac{a}{k^s}\right) - \Delta^2\left(\frac{b}{k^{s+1}}\right)\right]\left[\Delta^2\left(\frac{b}{k^{s+1}}\right)\frac{ms}{s+2}\right]^{-1} < \frac{\varepsilon_2}{k^\gamma} + O\left(\frac{1}{k}\right)\,. \tag{20}$$

Proof. To prove this theorem we transform the difference Emden-Fowler equation (1) to the system of difference equation

$$\Delta Y_0(k) = F_1(k, Y_0(k), Y_1(k)) := \left(-\frac{s+1}{k} + O\left(\frac{1}{k^2}\right)\right)(-Y_0(k) + Y_1(k))\,, \tag{21}$$

$$\Delta Y_1(k) = F_2(k, Y_0(k), Y_1(k)) := \left(-\frac{s+2}{k} + O\left(\frac{1}{k^2}\right)\right)\left(\frac{ms}{s+2}Y_0(k) - Y_1(k) + O\left(\frac{1}{k}\right)\right) \tag{22}$$

using the following change of variables

$$u(k) = \frac{a}{k^s} + \frac{b}{k^{s+1}}(1 + Y_0(k)), \tag{23}$$

$$\Delta u(k) = \Delta\left(\frac{a}{k^s}\right) + \Delta\left(\frac{b}{k^{s+1}}\right)(1 + Y_1(k)), \tag{24}$$

$$\Delta^2 u(k) = \Delta^2\left(\frac{a}{k^s}\right) + \Delta^2\left(\frac{b}{k^{s+1}}\right)(1 + Y_2(k)) \tag{25}$$

where $Y_i(k)$, $i = 0, 1, 2$ are new unknown functions. For computational details how the system (21), (22) is derived by transformation (23)–(25) we refer to [4, Part 3]). Consider functions b_1, b_2, c_1 and c_2, defined as follows:

$$b_1(k) := -\frac{\varepsilon_1}{k^\gamma}, \quad c_1(k) := \frac{\varepsilon_2}{k^\gamma}, \quad b_2(k) := -\frac{\varepsilon_3}{k^\beta}, \quad c_2(k) := \frac{\varepsilon_4}{k^\beta}$$

where $\varepsilon_i > 0$, $i = 1, 2, 3, 4$, $\beta > 0$ and $\gamma > 0$. These functions satisfy inequalities (8). Then

$$B_1(k, Y) := -Y_0 + b_1(k) = -Y_0 - \frac{\varepsilon_1}{k^\gamma}, \quad C_1(k, Y) := Y_0 - c_1(k) = Y_0 - \frac{\varepsilon_2}{k^\gamma}$$

and

$$B_2(k, Y) := -Y_1 + b_2(k) = -Y_1 - \frac{\varepsilon_3}{k^\beta}, \quad C_2(k, Y) := Y_1 - c_2(k) = Y_1 - \frac{\varepsilon_4}{k^\beta}\,.$$

To apply Lemma 1, the following inequalities must be valid:

$$F_1(k, b_1(k), Y_1) < b_1(k+1) - b_1(k) \tag{26}$$

if $(k, Y_0, Y_1) \in \Omega_B^1$ where

$$\Omega_B^1 := \left\{(k, Y_0, Y_1)\colon k \in \mathbb{N}(k_0), Y_0 = -\frac{\varepsilon_1}{k^\gamma},\ -\frac{\varepsilon_3}{k^\beta} \le Y_1 \le \frac{\varepsilon_4}{k^\beta}\right\},$$

(we refer to (9) where $i = 1$),

$$F_1(k, c_1(k), Y_1) > c_1(k+1) - c_1(k), \tag{27}$$

if $(k, Y_0, Y_1) \in \Omega_C^1$ where

$$\Omega_C^1 := \left\{(k, Y_0, Y_1)\colon k \in \mathbb{N}(k_0), Y_0 = \frac{\varepsilon_2}{k^\gamma},\ -\frac{\varepsilon_3}{k^\beta} \le Y_1 \le \frac{\varepsilon_4}{k^\beta}\right\},$$

(we refer to (10) where $i = 1$),

$$F_2(k, Y_0, b_2(k)) < b_2(k+1) - b_2(k), \tag{28}$$

if $(k, Y_0, Y_1) \in \Omega_B^2$ where

$$\Omega_B^2 := \left\{(k, Y_0, Y_1)\colon k \in \mathbb{N}(k_0), -\frac{\varepsilon_1}{k^\gamma} \le Y_0 \le \frac{\varepsilon_2}{k^\gamma},\ Y_1 = -\frac{\varepsilon_3}{k^\beta}\right\},$$

(we refer to (9) where $i = 2$), and

$$F_2(k, Y_0, c_2(k)) > c_2(k+1) - c_2(k) \tag{29}$$

if $(k, Y_0, Y_1) \in \Omega_C^2$ where

$$\Omega_C^2 := \left\{(k, Y_0, Y_1)\colon k \in \mathbb{N}(k_0), -\frac{\varepsilon_1}{k^\gamma} \le Y_0 \le \frac{\varepsilon_2}{k^\gamma},\ Y_1 = \frac{\varepsilon_4}{k^\beta}\right\},$$

(we refer to (10) where $i = 2$). From assumptions (12) and (13) we have $ms > 0$ and $s + 1 > 0$. These inequalities are used tacitly below. Now, we will verify inequalities (26)–(29). Conditions for their validity are, due to assumptions (12), (13) different from those derived in [4]).

Let us verify inequality (26). It will hold if

$$\begin{aligned} F_1(k, b_1(k), Y_1) \le \max_{(k,Y_0,Y_1)\in\Omega_B^1} F_1(k, b_1(k), Y_1) &= \left(-\frac{s+1}{k} + O\left(\frac{1}{k^2}\right)\right)\cdot\left(\frac{\varepsilon_1}{k^\gamma} - \frac{\varepsilon_3}{k^\beta}\right) \\ &< b_1(k+1) - b_1(k) = \frac{\varepsilon_1\gamma}{k^{\gamma+1}}\left(1 + O\left(\frac{1}{k}\right)\right). \end{aligned}$$

This inequality will hold if either

$$\gamma < \beta \tag{30}$$

or

$$\gamma = \beta,\quad \varepsilon_3 < \varepsilon_1\frac{\gamma + s + 1}{s+1}\,. \tag{31}$$

Now, verify inequality (27). It will hold if

$$F_1(k, c_1(k), Y_1) \geq \min_{(k,Y_0,Y_1)\in\Omega_C^1} F_1(k, c_1(k), Y_1) = \left(-\frac{s+1}{k} + O\left(\frac{1}{k^2}\right)\right) \cdot \left(\frac{-\varepsilon_2}{k^\gamma} + \frac{\varepsilon_4}{k^\beta}\right)$$
$$> c_1(k+1) - c_1(k) = -\frac{\varepsilon_2\gamma}{k^{\gamma+1}}\left(1 + O\left(\frac{1}{k}\right)\right).$$

This inequality will hold if either

$$\gamma > \beta \tag{32}$$

or

$$\gamma = \beta, \quad \varepsilon_4 < \varepsilon_2 \frac{\gamma + s + 1}{s+1}. \tag{33}$$

Let us verify inequality (28). It will hold if

$$F_2(k, Y_0, b_2(k)) \leq \max_{(k,Y_0,Y_1)\in\Omega_B^2} F_2(k, Y_0, b_2(k))$$
$$= \left(-\frac{s+2}{k} + O\left(\frac{1}{k^2}\right)\right)\left(\frac{ms}{s+2}\frac{-\varepsilon_1}{k^\gamma} + \frac{\varepsilon_3}{k^\beta} + O\left(\frac{1}{k}\right)\right)$$
$$< b_2(k+1) - b_2(k) = \frac{\varepsilon_3\beta}{k^{\beta+1}}\left(1 + O\left(\frac{1}{k}\right)\right).$$

This inequality will hold if either

$$\gamma > \beta \tag{34}$$

or

$$\gamma = \beta, \quad \gamma < 1, \quad \varepsilon_1 < \varepsilon_3 \frac{\gamma + s + 2}{ms}. \tag{35}$$

Note again that (34) contradicts to (30). Now, verify inequality (29). It will hold if

$$F_2(k, Y_0, c_2(k)) \geq \min_{(k,Y_0,Y_1)\in\Omega_C^2} F_2(k, Y_0, c_2)$$
$$= \left(-\frac{s+2}{k} + O\left(\frac{1}{k^2}\right)\right)\left(\frac{ms}{s+2}\frac{\varepsilon_2}{k^\gamma} - \frac{\varepsilon_4}{k^\beta} + O\left(\frac{1}{k}\right)\right)$$
$$> c_2(k+1) - c_2(k) = -\frac{\varepsilon_4\beta}{k^{\beta+1}}\left(1 + O\left(\frac{1}{k}\right)\right).$$

This inequality will hold if either

$$\gamma > \beta \tag{36}$$

or

$$\gamma = \beta, \quad \gamma < 1, \quad \varepsilon_2 < \varepsilon_4 \frac{\gamma + s + 2}{ms}. \tag{37}$$

Summing up all restrictions (30)–(37) we get the conditions (14)–(17). Inequalities (18)–(20) follow from inequalities (11). This concludes the proof of theorem. □

Remark 2. *The system of conditions* (14)–(17) *is the same as in the recent contribution [5]. But unlike of the paper [5] the range of admissible values of* m *and* s *is substantially enlarged. Hence, these new conditions we can use to formulate the following corollaries.*

3 COROLLARIES

In this part we derive two corollaries implied by Theorem 1.

Corollary 1. *Let either* (12) *or* (13) *hold. If, moreover,*

$$ms < \frac{(s+2)(s+3)}{s+1}\,, \tag{38}$$

then Theorem 1 is applicable.

Proof. If all hypotheses of the corollary hold, then inequality (38) implies the existence of ε_i, $i = 1, 2, 3, 4$ such that inequalities (14)–(17) hold. □

Remark 3. *The inequality* (38) *formally almost coincides with [5, inequality (4)]. But these inequalities are not equivalent because the mentioned inequality needs $s > 0$ and $m > 0$. This is not true in the case* (13)*.*

An advantage of the next corollary is that it makes it possible to apply Theorem 1 even if only some of the inequalities in terms of constants α and m, i.e., the constants involved in equation (1) hold. Additional conditions restricting the value of s are not necessary.

Corollary 2. *Let at least one of following assumptions* (i)–(iv) *hold:*

(i)

$$m \in \left(-7-4\sqrt{3}, -7+4\sqrt{3}\right), \quad -2 < \alpha < -m-1,$$

(ii)

$$0 < m < 1, \quad \alpha < -2,$$

(iii)

$$m > 1, \quad -2 < \alpha < \frac{1}{2}\left(-(m-1) + \sqrt{(m-1)^2+16m}\right),$$

(iv)

$$-2 < \alpha < -m-1, \quad m < 0, \quad (m-1)^2 + 16m > 0$$

and either

$$\alpha < \frac{1}{2}\left(-(m-1) - \sqrt{(m-1)^2+16m}\right)$$

or

$$\alpha > \frac{1}{2}\left(-(m-1) + \sqrt{(m-1)^2+16m}\right).$$

Then Corollary 1 is applicable.

Proof. It is easy to verify that, if one of assumptions (ii) or (iii) holds then so do inequalities (12). If assumptions (i) or (iv) hold, then so do inequalities (13).

Due to the conditions (12) or (13) we have $s + 1 > 0$ and inequality (38) in Corollary 1 can be written as

$$ms(s+1) < s^2 + 5s + 6.$$

This inequality was considered in [5] and implies [5, formula (16)]

$$(m-1)\left[\alpha^2+\alpha(m-1)-4m\right]<0. \tag{39}$$

If (12) holds then $m>0$. This case was treated in [5, Corollary 1] and leads to inequalities formulated in cases (ii) and (iii) in Corollary 2. So, we omit this part of considerations.
If (13) holds, then $m<0$ and $-1<s<0$. Using formula (5), these two inequalities imply

$$-2<\alpha<-m-1.$$

Therefore

$$m-1<0$$

and the inequality (39) is equivalent to the following one:

$$\alpha^2+\alpha(m-1)-4m>0. \tag{40}$$

Consider discriminant D of the quadratic equation

$$\alpha^2+\alpha(m-1)-4m=0 \tag{41}$$

with an unknown value of α. We have

$$D=(m-1)^2+16m.$$

If $D<0$, i.e., if

$$m\in(-7-4\sqrt{3},-7+4\sqrt{3})$$

then the inequality (40) holds for all α and inequalities in (i) are proved. If $D>0$ then the real distincts roots of quadratic equation (41) are

$$\alpha_{\pm}=\frac{1}{2}\left(-(m-1)\pm\sqrt{(m-1)^2+16m}\right)$$

and inequality (40) will hold if either

$$\alpha<\alpha_{-}=\frac{1}{2}\left(-(m-1)-\sqrt{(m-1)^2+16m}\right)$$

or

$$\alpha>\alpha_{+}=\frac{1}{2}\left(-(m-1)+\sqrt{(m-1)^2+16m}\right).$$

Thus, inequalities in (iv) are correct. □

Remark 4. *The result of Corollary 2 can be visualized in (m,α)-plane. The set of all points (m,α) satisfying at least one of assumptions (i)–(iv) is depicted in Figure 1. All such admissible points fill the yellow coloured open domain. This domain is bounded by the lines $m=1$, $\alpha=-2$, coloured in green, by the function*

$$\alpha=\frac{1}{2}\left(-(m-1)-\sqrt{(m-1)^2+16m}\right),$$

coloured in red and by the function

$$\alpha=\frac{1}{2}\left(-(m-1)+\sqrt{(m-1)^2+16m}\right),$$

coloured in blue. The subdomains I–IV depict points satisfying inequalities in (i)-(iv), respectively. A detail of a part of the domain IV is visualized in Figure 2.

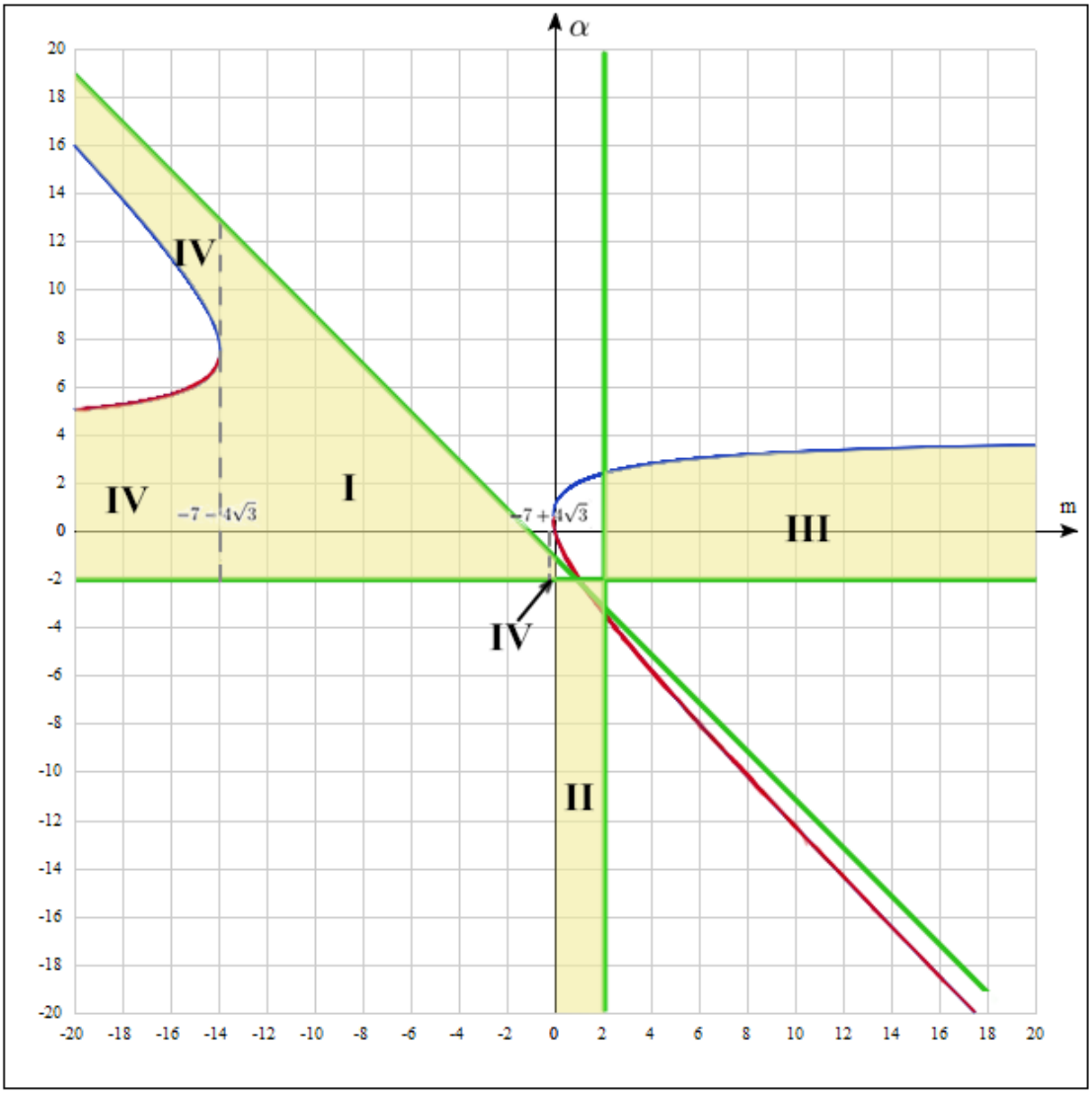


Figure 1: Admissible points of (m, α)-plane given by Corollary 2

4 EXAMPLE

Let $m = -4$, $\alpha = 1$. Then equation (1) reduces to the following

$$\Delta^2 u(k) \pm k \cdot u^{-4}(k) = 0. \tag{42}$$

Using (5) we get values s, a and b

$$s = -0.6, \quad a = \pm\sqrt[5]{\frac{25}{6}}, \quad b = \pm\frac{21}{25}\sqrt[5]{\frac{25}{6}}.$$

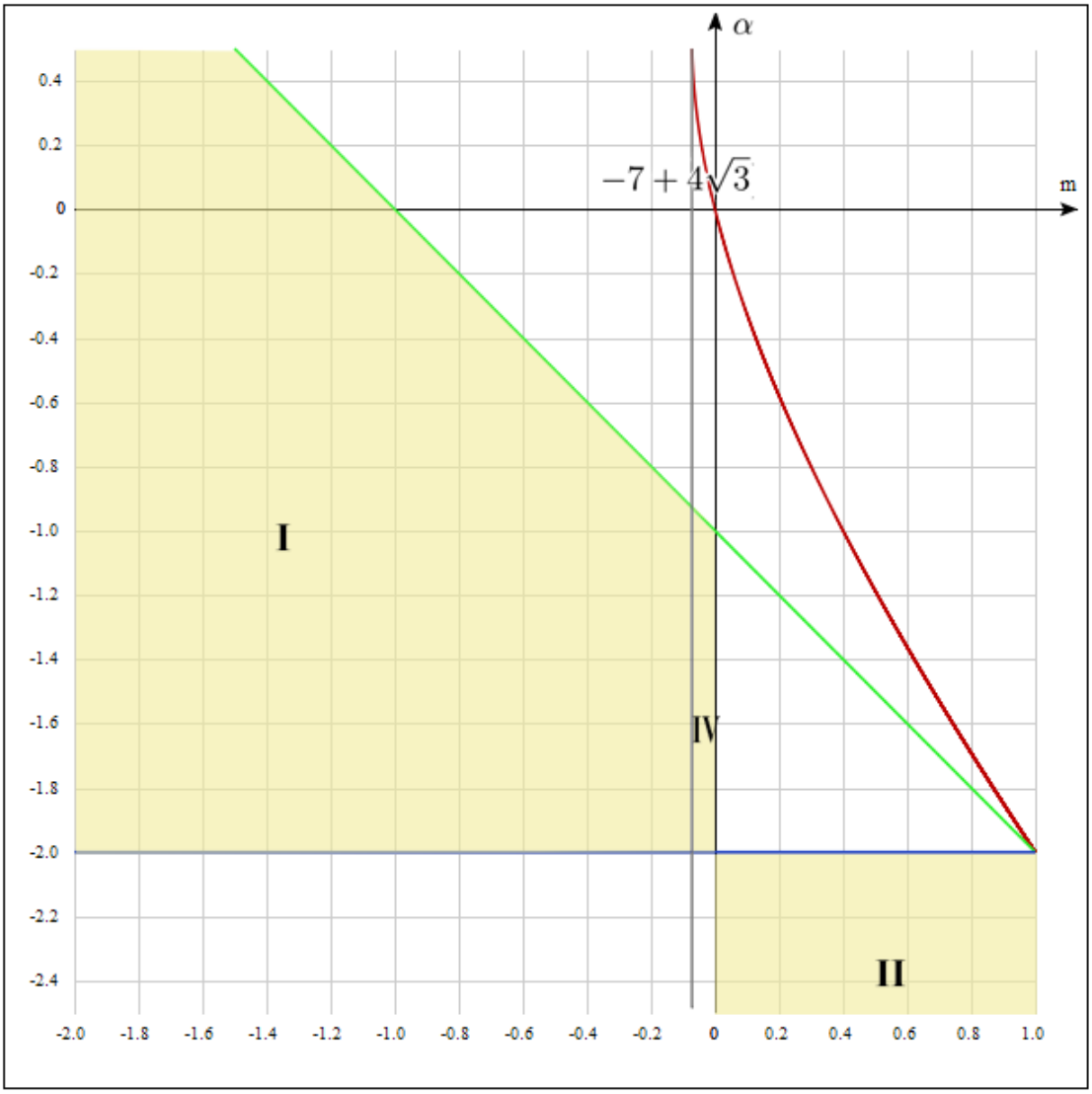


Figure 2: A detail of a part of the domain IV

Let $\varepsilon_1 = \varepsilon_2 = 1/2$, $\varepsilon_3 = \varepsilon_4 = 1$ and $\gamma = 3/4$. Then inequalities (14)–(17) reduce to

$$\varepsilon_3 = 1 < \varepsilon_1 \cdot \frac{\gamma + s + 1}{s + 1} = \frac{23}{16},$$

$$\varepsilon_4 = 1 < \varepsilon_2 \cdot \frac{\gamma + s + 1}{s + 1} = \frac{23}{16},$$

$$\varepsilon_1 = \frac{1}{2} < \varepsilon_3 \cdot \frac{\gamma + s + 2}{ms} = \frac{43}{48},$$

$$\varepsilon_2 = \frac{1}{2} < \varepsilon_4 \cdot \frac{\gamma + s + 2}{ms} = \frac{43}{48},$$

hence they are valid. Since (13) holds as well, all assumptions of Theorem 1 are fulfilled. Its statement says that, for sufficiently large fixed $k_0 > 0$, there exists a solution $u\colon \mathbb{N}(k_0) \to \mathbb{R}$ of

equation (42) such that, for every $k \in \mathbb{N}(k_0)$, inequalities (18), (19), (20) are valid. Let us write them in detail.

Analysis of inequality (18). Since

$$u(k) - \frac{a}{k^s} - \frac{b}{k^{s+1}} = u(k) \mp \sqrt[5]{\frac{25}{6}} k^{3/5} \mp \frac{21}{25} \sqrt[5]{\frac{25}{6}} \frac{1}{k^{2/5}},$$

$$\left[\frac{b}{k^{s+1}}\right]^{-1} = \left[\pm \frac{21}{25} \sqrt[5]{\frac{25}{6}} \frac{1}{k^{2/5}}\right]^{-1} = \pm \frac{25}{21} \sqrt[5]{\frac{6}{25}} k^{2/5}$$

and

$$\frac{\varepsilon_{1,2}}{k^\gamma} = \frac{1}{2k^{3/4}},$$

the inequality (18) turns into

$$-\frac{1}{2k^{3/4}} < \left[u(k) \mp \sqrt[5]{\frac{25}{6}} k^{3/5} \mp \frac{21}{25} \sqrt[5]{\frac{25}{6}} \frac{1}{k^{2/5}}\right] \left[\pm \frac{25}{21} \sqrt[5]{\frac{6}{25}} k^{2/5}\right] < \frac{1}{2k^{3/4}}\,, \tag{43}$$

or, after a simplification, to

$$\left| u(k) + \sqrt[5]{\frac{25}{6}} k^{3/5} + \frac{21}{25} \sqrt[5]{\frac{25}{6}} \frac{1}{k^{2/5}} \right| < \frac{21}{50} \sqrt[5]{\frac{25}{6}} \frac{1}{k^{23/20}}$$

in the case of equation (42) where the sign $+$ is applied and to

$$\left| u(k) - \sqrt[5]{\frac{25}{6}} k^{3/5} - \frac{21}{25} \sqrt[5]{\frac{25}{6}} \frac{1}{k^{2/5}} \right| < \frac{21}{50} \sqrt[5]{\frac{25}{6}} \frac{1}{k^{23/20}}$$

in the case of equation (42) where the sign $-$ is applied.

Analysis of inequality (19). Inequality (19) turns into

$$-\frac{1}{k^{3/4}} < \left[\Delta u(k) - \Delta\left(\frac{\pm \sqrt[5]{\frac{25}{6}}}{k^{-3/5}}\right) - \Delta\left(\frac{\pm \frac{21}{25} \sqrt[5]{\frac{25}{6}}}{k^{2/5}}\right)\right] \left[\Delta\left(\frac{\pm \frac{21}{25} \sqrt[5]{\frac{25}{6}}}{k^{2/5}}\right)\right]^{-1} < \frac{1}{k^{3/4}} \tag{44}$$

or, after a simplification, to

$$\left| \Delta u(k) - \sqrt[5]{\frac{25}{6}} \Delta\left(k^{3/5}\right) - \frac{21}{25} \sqrt[5]{\frac{25}{6}} \Delta\left(\frac{1}{k^{2/5}}\right) \right| < -\frac{21}{25} \sqrt[5]{\frac{25}{6}} \frac{\Delta(k^{-2/5})}{k^{3/4}}$$

in the case of equation (42) where the sign $+$ is applied and to

$$\left| \Delta u(k) + \sqrt[5]{\frac{25}{6}} \Delta\left(k^{3/5}\right) + \frac{21}{25} \sqrt[5]{\frac{25}{6}} \Delta\left(\frac{1}{k^{2/5}}\right) \right| < -\frac{21}{25} \sqrt[5]{\frac{25}{6}} \frac{\Delta(k^{-2/5})}{k^{3/4}}$$

in the case of equation (42) where the sign $-$ is applied.

Analysis of inequality (20). Inequality (20) turns into

$$-\frac{1}{2k^{3/4}}+O\left(\frac{1}{k}\right)$$
$$<\left[\Delta^2u(k)-\Delta^2\left(\frac{\pm\sqrt[5]{\frac{25}{6}}}{k^{-3/5}}\right)-\Delta^2\left(\frac{\pm\frac{21}{25}\sqrt[5]{\frac{25}{6}}}{k^{2/5}}\right)\right]\left[\Delta^2\left(\frac{\pm\frac{21}{25}\sqrt[5]{\frac{25}{6}}}{k^{2/5}}\right)\frac{12}{7}\right]^{-1}$$
$$<\frac{1}{2k^{3/4}}+O\left(\frac{1}{k}\right). \quad (45)$$

or, after a simplification, to

$$\left|\Delta^2u(k)-\sqrt[5]{\frac{25}{6}}\Delta^2\left(k^{3/5}\right)-\frac{21}{25}\sqrt[5]{\frac{25}{6}}\Delta^2\left(\frac{1}{k^{2/5}}\right)\right|$$
$$<\frac{36}{25}\sqrt[5]{\frac{25}{6}}\left(\frac{1}{2k^{3/4}}+O\left(\frac{1}{k}\right)\right)\Delta^2(k^{-2/5})$$

in the case of equation (42) where the sign $+$ is applied and to

$$\left|\Delta^2u(k)+\sqrt[5]{\frac{25}{6}}\Delta^2\left(k^{3/5}\right)+\frac{21}{25}\sqrt[5]{\frac{25}{6}}\Delta^2\left(\frac{1}{k^{2/5}}\right)\right|$$
$$<\frac{36}{25}\sqrt[5]{\frac{25}{6}}\left(\frac{1}{2k^{3/4}}+O\left(\frac{1}{k}\right)\right)\Delta^2(k^{-2/5})$$

in the case of equation (42) where the sign $-$ is applied.

Remark 5. *Inequality* (38) *is also fulfilled because*

$$\frac{12}{5}=ms<\frac{(s+2)(s+3)}{s+1}=\frac{84}{10}.$$

Then Corollary 1 is applicable as well.

CONCLUSION

In the paper we generalized the results on asymptotic behaviour of solutions of Emden-Fowler type difference equation (1) derived in [5]. The progress was achieved by a new estimation of the right-hand sides of the auxiliary system (21), (22) if a new set of assumptions on coefficients of equation (1) is used (inequalities (13) in Theorem 1). This generalization is clearly visible in Figure 1 where the sets II and III of admissible points (m,α) are implied by the results of the paper [5] and the sets I and IV are new domains of admissible points not covered by the results of [5].

Let us note, referring to [1], that the classic second-order differential Emden-Fowler equation

$$y''(x)\pm x^{\alpha}y^{m}(x)=0$$

has the exact solution

$$y = a/x^s,$$

where the coefficients a and s are computed by formulas (5). The formula (4) describes the asymptotic behaviour of a solution to (1) being similar to the behaviour of this exact solution. Therefore it seems, if in (1) the difference $\Delta u(k)$ is redefined as a difference with the step equaling an arbitrarily small positive number h rather than with one equalling 1 that is, if

$$\Delta u(k) := \left(u(k+h) - u(k)\right)/h,$$

that, for $h \to 0$, the yellow domain on Figure 1 should cover almost all (m, α)-plane (except for the value $m = 1$ and maybe values $\alpha = 0$, $m = 0$). This is still an open problem.

Moreover, the method of investigation used seems to be suitable for analyzing the asymptotic behaviour of generalized Emden-Fowler type difference equations such as

$$\Delta^2 u(k) + c_1 k_1^\alpha u^{m_1}(k) + c_2 k_2^\alpha u^{m_2}(k) + \cdots + c_\ell k_\ell^\alpha u^{m_\ell}(k) = 0$$

where c_i, α_i and m_i, $i = 1, \ldots, \ell$ are suitable constants.

References

[1] Bellman, R., *Stability Theory in Differential Equations.* Dover Publications, Inc., New York, 2008, 176 pp.

[2] Diblík, J., Discrete retract principle for systems of discrete equations. *Comput. Math. Appl.*, 42, 2001, 515–528 .

[3] Diblík, J., Asymptotic behavior of solutions of discrete equations. *Funct. Differ. Equ.*, No 11(1–2), 2004, 37–48.

[4] Diblík, J., Korobko, E., Solutions of perturbed second-order discrete Emden-Fowler type equation with power asymptotics of solutions. *Mathematics, Information Technologies and Applied Sciences 2020, post-conference proceedings of extended versions of selected papers*, Brno: UNOB Brno, 2020, 30–44. ISBN: 978-80-7582-366-3.

[5] Korobko, E., Asymptotic characterization of solutions of Emden-Fowler type difference equation. *The Student conference EEICT 2021, Faculty of Electrical Engineering and Communication, Proceedings II of the 27th Conference STUDENT EEICT 2021, Selected Papers*, Brno University of Technology, 2021, 250–254.

Acknowledgement

This research has been supported by the project of specific university research at Brno University of Technology, Faculty of Electrical Engineering and Communication, FEKT-S-20-6225.

CONTRIBUTION OF PREPARATORY MATH COURSE FOR FIRST-YEAR UNIVERSITY STUDENTS: BAYESIAN APPROACH

Petr Emanovský
Palacky University Olomouc
17. listopadu 12, petr.emanovsky@upol,cz

Abstract: *Preparing university students for a mathematical test can take place in a variety of ways. One option is to take a suitable praparatory course, which is often offered mainly to first-year students. The natural question then arises whether completing such a course has a significant effect on students' success in the test. The research described in this paper is focused on the relationship between attending the special preparatory course for prospective teachers and their success in the mathematical test. Two ways of processing data from contingency table relating student test results and kind of their training (course, individual) are shown. First one is the traditional independence testing based on the Pearson chi-squared statistic and the second one is the Bayesian model comparing. In both cases, the research results indicate a statistically significant difference in test score in favor of students attended the course. In addition, these results are supported by direct calculation of posterior conditional probabilities of student's attending the course under the condition of successful test passing. Using the Bayes rule it was shown that the probability that a randomly selected successful student attended the course is greater than 80%.*

Key words: contingency table, preparatory course, university teacher training, Bayesian approach.

INTRODUCTION

The success or failure of students in exams, especially during the first year of study, has a great influence on their motivation for further study. This motivation is especially significant for prospective teachers, as it can be reflected in their future pedagogical work ([1], [19]). For this reason, a number of optional preparatory courses are used to offer a teacher training to students at faculties implementing teacher education. The preparatory courses play an important role especially for pre-service mathematics teachers, as a certain level of basic knowledge is necessary for the study of mathematics. A number of studies showed the importance of quality training of future mathematics teacher for their beliefs and attitudes towards mathematics ([18], [25], [27]). On the other hand, not many pre-service teachers programs have been investigated with respect to their effectiveness and influence on test success ([15]). Obviously, some initial dispositions of students play an important role for successful study ([2]). Especially preceding domain knowledge and previous learning experiences represent the significant factors influencing the success ([10]). Of course, it is not possible to forget about other important factors, such as student motivation ([1], [7], [19]), curricula design ([24]), cultural expectation ([22]), etc.

Students of mathematics teacher training at the Faculty of Science of Palacky University in Olomouc, Czech Republic take an algebra test during the first year. The test is primarily focused on algebraic symbol manipulation skills and basic knowledge of linear algebra. The highest rating of test points is 100 and at least 60 points are considered a successful result. The success of students in the test can be influenced by many different factors. An important prerequisite for success is, of course, the active preparation of students in solving the

recommended tasks. Students can solve these tasks individually or as part of an optional preparatory course under the guidance of a teacher. The preparatory course is focused on the acquisition of basic algebraic concepts and their active use in problem solving.

1 MATERIALS AND METHODS

1.1 Research Problem and Research Hypotheses

The research problem was to find out the relationship between attending the preparatory course for first-year students of future teacher training and their success in the mathematical test. The main aim of this research was to answer the question whether the preparatory course contributed significantly to the successful writing of the final mathematical test. With regard to this question, the null and alternative hypotheses were formulated as follows:

Hypothesis H_0: The difference between test results of students completing the preparatory course and other students is not statistically significant.
Hypothesis H_A: The difference between test results of students completing the preparatory course and other students is statistically significant.

1.2 General Background

Solving the research problem supposes to carry out an independence test for two categorical variables – the results of students' mathematical test and the kind of students' preparation for the test. The frequentist approach of the independence testing is based on the familiar Pearson chi-squared statistics. In this test, observed data are compared with the expected ones under an independence model. The dependence between the variables is then determined by the statistical significance of the difference. The Bayesian perspective is completely different. This viewpoint compares two possible models – the model M_K for dependent variables and model M_I for independent variables.

1.3 Participants

The sample consisted of 116 first year mathematics teacher training students at Faculty of Science of Palacky University in Olomouc. The mathematical test was administered to these students in 2020 at the end of the first semester. 74 students (63.8 %) completed the preparatory course and 42 students (36.2 %) prepared individually (Tab. 1).

1.4 Research Instrument

According to the research problem comparative research design was adopted. Data on the two categorical variables obtained from the sample are presented using the contingency table (Tab. 1). The variable "kind of students' preparation" takes two values – "course" and "individual". The values of the variable "results of test" are "successful" (score 60 – 100) and "unsuccessful" (score 0 – 59).

	Course	**individual**	**Σ**
Successful	49	19	68
Unsuccessful	25	23	48
Σ	74	42	116

Tab. 1: Contingency table relating student test results and kind of their training

1.5 Data Analysis and Results

Program R with its package "*LearnBayes*" can be used for the data analysis ([3]). The following sequence of R language commands creates the contingency table as variable "*student*". Using R function *chisq.test* to test the independence hypothesis one obtains *p*-value approximately 0.04, which is an evidence that the result of the test is related to the completion of the course.

```
> student=matrix(c(49,25,19,23),c(2,2))
> student

     [,1] [,2]
[1,]   49   19
[2,]   25   23

> chisq.test(student)

        Pearson's Chi-squared test with Yates' continuity correction

data:  student
X-squared = 4.0346, df = 1, p-value = 0.04458
```

Testing the independence of two categorical variables can be also performed using Bayesian model comparison. The easiest way is to compare our model "*student*" with the independent model "*uniform*" using R function *ctable*. The Bayes factor approximately 3.5 > 1 indicates support against independence of the variables.

```
> uniform=matrix(rep(1,4),c(2,2))
> uniform

     [,1] [,2]
[1,]    1    1
[2,]    1    1

> ctable(student,uniform)

[1] 3.51763
```

The Bayesian approach makes it possible to compare an "independence model" with a model close to an "independence" one. Such a model has been introduced by Albert and Gupta in [4]. This model is based on conjugated Dirichlet distribution and allows the calculation of Bayes factors for various models approaching the „independence" one. The following short algorithm computes the Bayes factors to compare our model „student" with an alternative models that are close to independence.

```
> log.k=seq(2,8)
> comp.log.bf=function(log.k)
+ log(bfindep(student,exp(log.k),1000000)$bf)
> log.bf=sapply(log.k,comp.log.bf)
> bf=exp(log.bf)
> round(data.frame(log.k,log.bf,bf),2)
```

log.k	log.bf	bf
1	2	-0.10 0.91
2	3	0.81 2.25
3	4	1.01 2.75
4	5	0.76 2.14
5	6	0.41 1.51
6	7	0.18 1.20
7	8	0.07 1.07

The key function *bfindep* required the Dirichlet parameter k and the size of the simulated sample (1000000 in the case). The output is a list of Bayes factors (bf) for the sequence of the values of log k. The maximum value 2.75 indicates some support for the model with log $k = 4$ that is close to the independence model against the independence one (Fig. 1). All the results of the data analysis confirm the positive effect of completing the preparatory course on the students‘ success in the test.

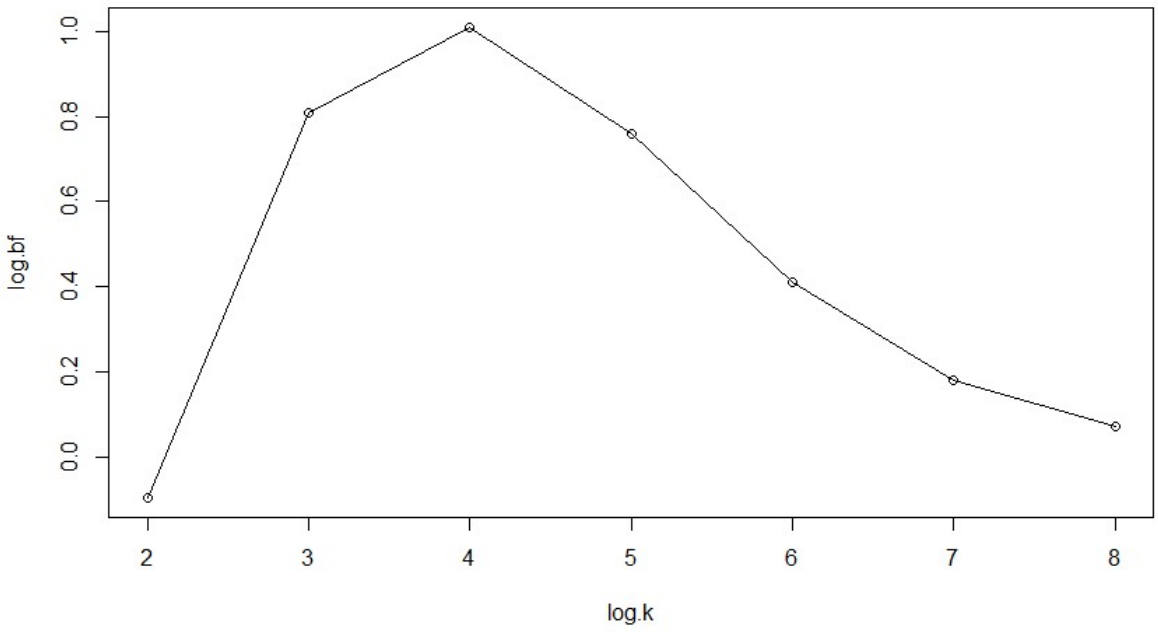


Fig. 1: Log Bayes factor in support of model M_K over M_I (source: own calculation)

2 USEFULNESS OF BAYESIAN RULE

The posterior conditional probability that a randomly selected successful student attended the course should be determined as a secondary „research problem“. For this purpose, let us consider the following statements:
Hypothesis H_1: Randomly selected student completed the preparatory course.
Hypothesis H_2: Randomly selected student did not complet the preparatory course.

Proposition T: Randomly selected student has written the test successfully.
Probability of the hypothesis H_1, $P(H_1) = \frac{74}{116} \approx 0.638$. For probability of the hypothesis H_2 it holds $P(H_2) = \frac{42}{116} \approx 0.362$. From previous records, it follows for the conditional probabilities $P(T/H_1) = 0.75$ and $P(T/H_2) = 0.3$.
Hence, $P(T) = P(H_1)\,P(T/H_1) + P(H_2)\,P(T/H_2) = 0.638 \cdot 0.75 + 0.362 \cdot 0.3 = 0.588$.
Thus, almost 60% of students take the test for a long time.
However, we are interested in a specific student who was successful in the test. The prior probability 0.638 can be modified after finding her/his success in the test to the posterior probability using the Bayes rule
$P(H_1/T) = \frac{P(H_1)\,P(T/H_1)}{P(T)} = 0.815$.
This means, that the probability that a randomly selected successful student attended the course is greater than 80 %. Consequently, the complementary probability is 0.185.
$P(H_2/T) = \frac{P(H_2)\,P(T/H_2)}{P(T)} = 0.185$.
Thus, a randomly selected successful student prepared for the test individually with a probability of 18.5 %.

These facts also speak in favor of organized teacher-led training.

DISCUSSION

The research results described above clearly showed the usefulness of the preparatory course for future teachers. The findings of this research correspond to the results of similar studies. For instance, [9] describe the effect of the preparatory online course on the success of study of chemistry. The positive correlation between the examination results in mathematics and the attendance of a preparatory course for first-year university students focused on the basic mathematics skills was confirmed in the study [17]. Other studies showing the positive effect of the courses are [5], [13], [14]. On the other hand, it is fair to mention also the studies that have found no significant effect of the preparatory courses ([11], [23]). These inconsistent results can be caused by different type of the courses. There exist a lot of studies focused on the effect of the preparatory courses but only a few of them investigate this effect in relation to the type of a course ([6], [16]). The importance of online learning in the context of the global coronavirus pandemic is becoming increasingly important. Further research on pending development of the preparatory courses and evaluation of preparatory material is needed, in particular in the field of e-learning, with regard to support for the independent learning ([17]).
It is a debatable point whether the test performance is the best manner to investigate the educational benefit of a course. It would be desirable to have not only the results of a compulsory math test, but also more information about students' future performance. On the other hand, this is an easily accessible and universally applicable measure of achievement.
Investigating the dependence of categorical data in a contengency table using Bayesian factors brings some specific benefits. For example, according to a Bayes factor one can quantify evidence in behalf of the null hypothesis and the factors can be monitored as the data acumulate ([28]). The second advantage can be particularly important when the data come from a natural process that develops over time without any predetermined stopping point ([20]). It can be argued that these benefits are well known for a long time, as Bayes factors for contingency tables have been introduced more than fifty years ago ([21]). However, most researches usually use for the analysis of contingency tables classical methods, obtaining *p*-values through Pearson chi-square statistics or likelihood ratio tests. One of the reasons for the omission of Bayesian methods in the empirical research was the lack of their implementation

in user-friendly software packages. Recently, however, the situation in this area has greatly improved with the appearance of new computer support.

CONCLUSION

Correct understanding of basic mathematical concepts and thorough acquisition of numerical skills is a key prerequisite for successful study of all mathematical disciplines taught at the university. This study shows that a suitable preparatory course for beginning students focused on these basic knowledge and skills can contribute to the successful writing of the final mathematics test. It is natural to assume that this successful start will positively affect their entire further study.

The relatively less used Bayessian approach to processing data from contingency table is shown in the article. The possibility of a Bayesian approach was suggested in the data analysis section. This approach represents an alternative to classical data analysis. Classical (frequentist) statistical methods use probabilistic models applicable only to mass phenomena whose occurrence or absence can be observed repeatedly in many situations. Probability is understood here as the relative frequency of occurrence of the observed phenomenon ([26]). On the other hand, the Bayesian approach interprets probability as a measure of belief in the truth of the statement ([12]). Although this approach is historically older, it has been criticized for its subjectivism and has been outside the main scientific interest. Since the 1990s, however, the methods have experienced a certain renaissance and today are considered modern methods with wide practical application.

Any researcher who is serious about statistical data analysis should not be limited to the classical frequentist methods, but should also become familiar with some alternative approaches. The Bayesian perspective is in a sense more universal than the classical one, because Bayesian methods are not limited to the analysis of mass phenomena. Admittedly, these methods are more demanding on theoretical knowledge and numerical skills. However, Bayesian statistical analysis has recently been intensively developed, among other things, thanks to new algorithmic procedures and suitable freely accessible software (Python, WinBugs, R, Jasp).

References

[1] Afzal, H., Ali, I., Khan, M., A. & Hamid, K. A Study of University Students' Motivation and Its Relationship with Their Academic Performance. *International Journal of Business and Management*, 5(4), 2010, p. 80 – 88.

[2] Anthony, G. Factors influencing first-year students' success in mathematics. *International Journal of Mathematical Education in Science and Technology,* 31(1), 2000, p. 3 – 14. https://doi.org/10.1080/002073900287336

[3] Albert, J. *Bayesian Computation with R.* Springer, 2009.

[4] Albert, J. & Gupta, A. Mixtures of Dirichlet distributions and estimation in contingency tables. *Annals of Statistics,* 10, 1981, p. 1261-1268.

[5] Ballard, C. L., & Johnson, M. F. Basic math skills and performances in an introductory economics class. *Journal of Economic Education*, 35(1), 2004, p. 3–23.

[6] Biehler, R., Bruder, R., Hochmuth, R., Koepf, W., Bausch, I., Fischer, P. R. & Wassong, T. VEMINT – Interaktives Lernmaterial für mathematische Vor- und Brückenkurse. In I.

Bausch, R. Biehler, R. Bruder, P. R. Fischer, R. Hochmuth, W. Koepf, S. Schreiber, & T. Wassong (Eds.), *Mathematische Vor- und Brückenkurse: Konzepte, Probleme und Perspektiven*. Wiesbaden: Springer Spektrum, 2014, p. 261–276.

[7] Biggs, J. *Higher Education Research and Development,* 12, 1993, p.73 – 85.

[8] Bolstad, W. M. *Itroduction to Bayesian statistics.* John Wiley, 2007.

[9] Botch, B., Day, R., Vining, W., Stewart, B., Rath, K., Peterfreund, A. & Hart, D. Effects on Student Achievement in General Chemistry Following Participation in an Online Preparatory Course. *Journal of Chemical Education,* 84 (3), 2007, p. 547 – 553.

[10] Crowford, K., Gordon, S., Nicholas & Prosser, M. *Learning and Instruction.* 8, 1998, p. 255 – 468.

[11] Di Pietro, G. The short-term effectiveness of a remedial mathematics course: evidence from a UK university, *IZA Discussion Paper Series*, No. 6358, Bonn, 2012.

[12] Emanovský, P. Bayesian versus frequentist approach to statistical inference. In M. Hrubý & E. Kolářová (Eds.), *Mathematics, Information Technologies and Applied Sciences.* Brno: University of Defence, 2020. [Online]. [Cit. 2021-07-26]. Available at: <http://mitav.unob.cz/data/MITAV 2020 Proceedings.pdf>.

[13] Engelbrecht, J. C. Academic support in mathematics in a Third World environment. *Journal of Computers in Mathematics and Science Teaching*, 16(2), 1997, p. 323–333.

[14] Espey, M. Testing math competency in introductory economics. *Review of Agricultural Economics*, 19(2), 1997, p. 484–491.

[15] Fernandes, D. Analyzing four preservice teachers' knowledge and thoughts through their biographical histories. *Proceedings of the Nineteenth International Conference for the Psychology of Mathematics Education*, Vol. II, Universidade Federal de Pernambuco, Recife, Brazil, 1995, p. 162 – 169.

[16] Fischer, P. R. *Mathematische Vorkurse im Blended-Learning-Format.* Wiesbaden: Springer, 2014.

[17] Greefrath, G., Koepf, W. & Neugebauer, Ch. Is there a link between Preparatory Course Attendance and Academic Success? A Case Study of Degree Programmes in Electrical Engineering and Computer Science. *Int. J. Res. Undergrad. Math. Ed.*, 3, 2017, p. 143–167.

[18] Hancock, D. R., Bray, M. & Nason, S. A. Influencing University Students' Achievement and Motivation in a Technology Course. *The Journal of Educationa Research*, 95 (6), 2010, p. 365 – 372.

[19] Hill, A. P. Motivation and university experience in first-year university students: A self-determination theory perspective. *Journal of Hospitality, Leisure, Sport & Tourism Education,* 13, 2013, p. 244 – 254. https://doi.org/10.1016/j.jhlste.2012.07.001

[20] Jamil, T., Ly, A., Morey, R. D., Lovel, J., Marsman, M. & Wagenmakers, E. J. Default "Gunel and Dickey" Bayes factors for contingency tables. *Behavioral Research*, 49, 2017, p. 638 – 652.

[21] Jeffreys, H. Some tests of significance, treated by the theory of probability. *Proceedings of the Cambridge Philosophy Society*, 31, 1935, p. 203–222.

[22] Killen, R. *Higher Education Research and Development,* 13, 1994, p. 199 – 211.

[23] Lagerlöf, J. N. M., & Seltzer, A. J. The effects of remedial mathematics on the learning of economics: evidence from a natural experiment. *Journal of Economic Education*, 2009, p. 115–136.

[24] McKay, J. & Kember, D. *Higher Education Research and Development,* 16, 1997, p. 55 – 67.

[25] McLeod, D. B. Research on affect and mathematics learning in the JRME: 1970 to the present. *Journal for Research in Mathematics Education* 25 (6), 1994, p. 637–647.

[26] McMillan, J. H. & Schumacher, S. *Research in Education – Evidence-based Inquiry.* New Jersey: Pearson, 2010.

[27] Philippou, G. N. & Christou, C. The effects of a preparatory mathematics program in changing prospective teachers' attitudes towards mathematics. *Educational Studies in Mathematics,* 35, 1998, p. 189-206.

[28] Rouder, J. N. Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review*, 21, 2014, p. 301–308.

Acknowledgement

The work presented in this paper has been supported by the Palacky University project „Algebraic and Geometric Structures“ IGA PrF 2021 030

STATISTICS AS A SUPPORT FOR EXPERIMENTAL FINDINGS

Kamila Hasilová
Department of Quantitative Methods,
University of Defence, Brno, Czech Republic
kamila.hasilova@unob.cz

Milan Vágner
Department of Quantitative Methods,
University of Defence, Brno, Czech Republic
milan.vagner@unob.cz

Abstract: *In the article, we focus on one of possible ways how to increase the scientific quality of experimental theses. Experimental results are of great importance and conclusions drawn from them as well, but for the reader without the expert eye, it may be difficult to come to the same conclusion. Therefore, we apply an interdisciplinary approach with emphasis on the point of view of a statistician. On a particular example, we present how a thesis could be moved to a higher professional level with the help of a statistical approach to assessing experimental results.*

Keywords: experiment, assessment, statistics, functional data analysis, thesis, consultant.

INTRODUCTION

The highest level of university study is postgraduate study. An important part of it is the elaboration and subsequent defence of a dissertation thesis. The demands for the thesis are very high. Students are expected to demonstrate thorough knowledge of the current state of the problem and that they are able to work scientifically and independently. It is highly appreciated if the student, i.e. the author of the dissertation thesis, offers his own contribution or solution to the development of the scientific discipline. That is new, yet unpublished knowledge.

The students demonstrate the ability of a suitable selection of a professional problem and its processing in an appropriate manner. Their cooperation with scientific institutes is also expected since the scientific level of the theses is guaranteed by people who participate in the dissertation creation process, namely by supervisors and their professional erudition. In addition to the supervisor of the thesis, the author usually cooperates with a consultant. Depending on the nature of the work, consultations with other experts may be necessary. In some cases, a supervisor specialist may also be assigned. They all ensure that the work meets the relevant formal and content requirements, but especially that it brings appropriate results.

This is also the case at the University of Defence postgraduate study. The students select from the offered topics, which they work on under the guidance of experienced supervisors. The dissertation themes cover a wide range of specializations, depending on the study field. In accordance with the requirements for the dissertation scientific level, new findings, very often confirmed experimen-

tally, are incorporated into the theses. Students then draw appropriate conclusions based on the experimental results.

We pondered if these conclusions, which are usually based only on the visual expert assessment, could be supported by a suitable mathematical approach. More precisely, to use the means of statistics to confirm the experimental results and assessment, which can take the scientific quality of the dissertation thesis to a higher level. We understand that the aim of the thesis is not a statistical evaluation of the results, but a proposal to verify a certain procedure. Nevertheless, we believe that some statements could be assisted with an appropriate statistical test. This would eliminate possible incorrect conclusions or support the correct conclusions, respectively. Students could be assisted by a supervisor specialist dealing with statistics.

SPECIFIC EXAMPLE

In the following text, we introduce one of possible solutions how the statistics could improve the dissertation thesis to support the assessment of the experimental results. We selected thesis *Surface contamination control by using thermal desorption*, which belongs to the study field *Force and civil protection* and which was successfully defended in 2015[1]. The thesis is publicly available in the library of the University of Defence in Brno.

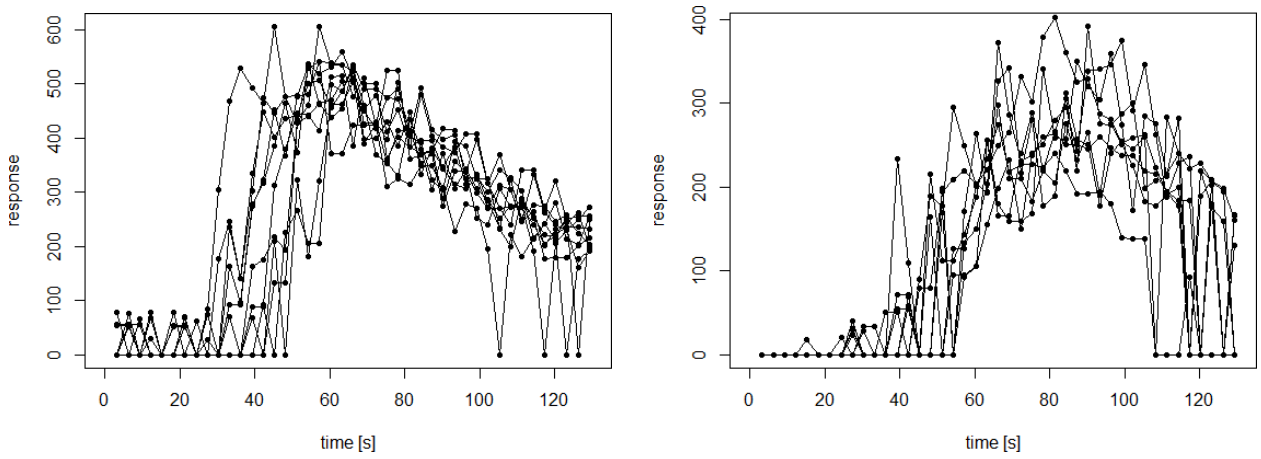
In the thesis, the possibility of using a decontamination chamber in combination with a portable automatic spectrometer is analysed and subsequently experimentally verified. The thesis deals with the control of the residual contamination (after decontamination) of surfaces contaminated by toxic chemical agents. Based on the experiments and their results, optimal conditions for the residual contamination control were proposed regarding the ambient temperature and the toxic agent involved. Consequently, a methodological procedure for carrying out efficient checks on decontamination in the chamber was established. The procedure can be used by the Integrated Rescue System.

The issue described above is certainly interesting and its processing is highly beneficial. From our professional point of view, it seems appropriate to support the experimental method, for example by testing, which would clearly confirm the accuracy of the experiment. Based on the data from the thesis, we are going to show how this work could be supplemented. For this purpose, we selected the measurements of nerve agent substitutes (denoted as substance A) and blister agent substitutes (denoted as substance B).

Let us start with one of the statements, which is connected with the data presented in Figure 1(a) [4, p. 77]: “The response is considered plausible only in the case of three repetitions at the same or higher level of response. In a measurement time interval from 30 to 102 seconds, the response of the detector varies within a certain response interval, which is considered sufficient to demonstrate the presence of the test substance.” Similarly for the other test substance (see Figure 1(b)), the

[1]The thesis was selected at random, it meets the required criteria for dissertation theses. We only want to show a possible way of cooperation between experimenters and consultants, not to speak against the thesis.

author of the thesis stated: “The response of the the detector lies in the time interval from 55 to 117 seconds.” [4, p. 79]



(a) substance A (b) substance B

Fig. 1. Measurements of the test substances

For the “untrained eye” of a reader, it might be difficult to see the mentioned intervals and the plausibility of the response. Therefore, we would like to propose statistical support for this type of assessment.

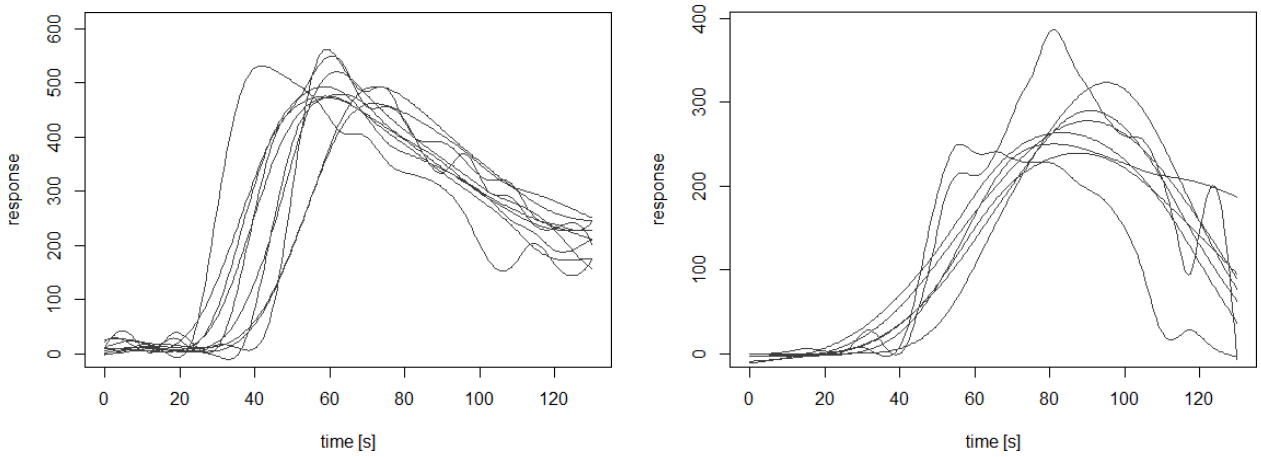
Numerical model

Concentration of a chemical substance is a continuous function of time. On the other hand, the measurements are usually carried out at discrete time points only. Therefore, we obtain discrete values, which represent a continuous function [8, 10].

The discretely measured data can be identified as a random sample of independent real-valued functions on a closed real interval [13]. Observations at the discrete points are influenced by measurement errors, which can be seen as a random variation around a smooth trajectory. Hence, the functional data analysis (FDA) provides an appropriate statistical way to model the discretely measured data [7, 14]. In the FDA framework, the observations are taken as curves reconstructed from the data, i.e. in simple words, *one curve = one observation*.

Several methods can be used for the reconstruction process, an individual function can be reconstructed using parametric, nonparametric or semiparametric methods, for details see, e.g. [10, 5, 11, 12]. Parametric methods offer a simple evaluation of the reconstructed function. However, applying polynomials, for example, brings a question how to select the order of the polynomial to capture the overall course of the data but not to oversmooth them [6]. On the other hand, one can use one of the nonparametric smoothing methods to reconstruct the individual curve, such as a moving average or the loess function [2].

For presentation purposes, we propose to apply smoothing splines (implemented in R [9]). The smoothing spline is a piecewise polynomial function with a certain order of smoothness. It balances a measure of goodness of fit of the function to the data and a measure of the smoothness (defined by derivatives of the function) [3]. The resulting smoothed curves are displayed in Figure 2.



(a) substance A (b) substance B

Fig. 2. Smoothed curves of discrete measurements from Figure 1

Having estimated the curves, we can proceed to a statistical test, which confirms or disproves the statements. The overall procedure is the same as in the univariate settings.

Statistical test

If we need to test the hypothesis that the mean function, $m(t)$, is equal to a prespecified function, $\mu_0(t)$, we use the one sample test. The null hypothesis takes the form $H\colon \mu(t) \equiv \mu_0(t)$ for $t \in [a, b]$ and the alternative $A\colon \mu(t) \neq \mu_0(t)$ for some $t \in [a, b]$. Test statistics is similar to the one from the univariate case

$$T = \frac{m(t) - \mu_0(t)}{v(t)} \cdot \sqrt{N},$$

where $v(t)$ stands for the functional standard deviation and N is the number of the observations, i.e. the number of curves. Under the null hypothesis, statistics T follows the Student t-distribution with $N - 1$ degrees of freedom [14].

Classic summary statistics, needed for the test, apply equally to functional data. Mean function $m(t)$ is the average of functions $f_j(t)$ $(j = 1, \ldots, N)$

$$m(t) = \frac{1}{N} \sum_{j=1}^{N} f_j(t),$$

variance function $v^2(t)$ is defined similarly

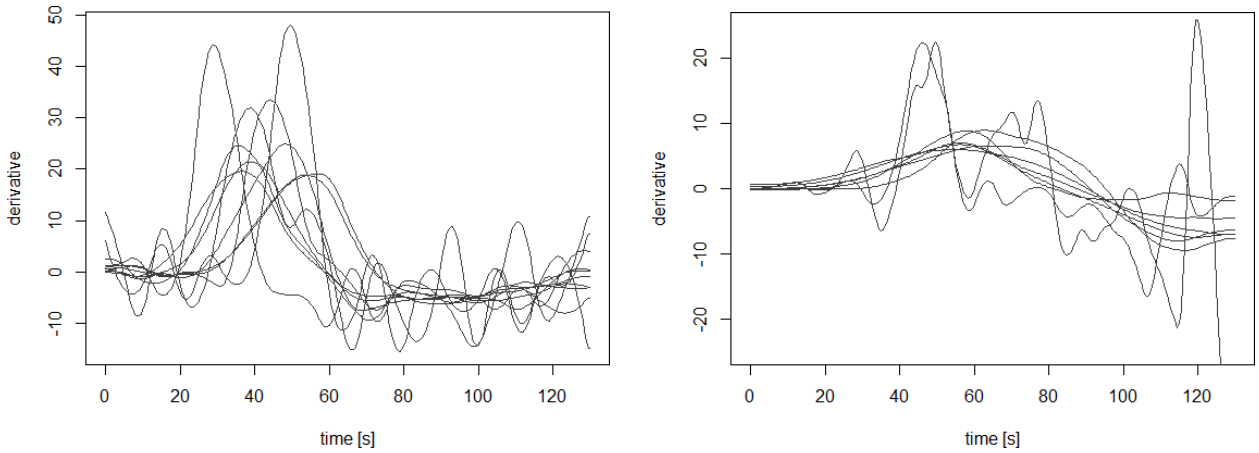
$$v^2(t) = \frac{1}{N-1}\sum_{j=1}^{N}[f_j(t) - m(t)]^2$$

and the standard deviation function is the square root of the variance function [10].

Statistical evaluation

Let us get back to the first part of the statement: "The response is considered plausible only in the case of three repetitions at the same or higher level of response." In other words, the function has to be nondecreasing. This is equivalent to the condition of the first derivative being positive. Having functional representation of the data, we can easily differentiate the curves and test the positivity of the derivatives.

The derivatives of the estimated curves are presented in Figure 3. The mean functions of the respective test substances with their confidence bands are shown in Figure 4.

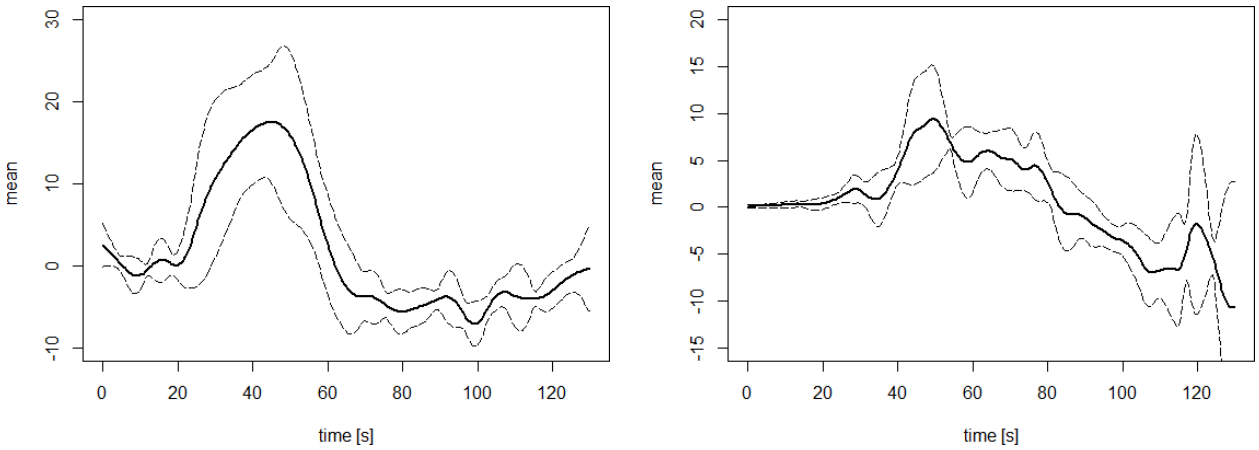


(a) substance A (b) substance B

Fig. 3. First derivatives of the smoothed curves

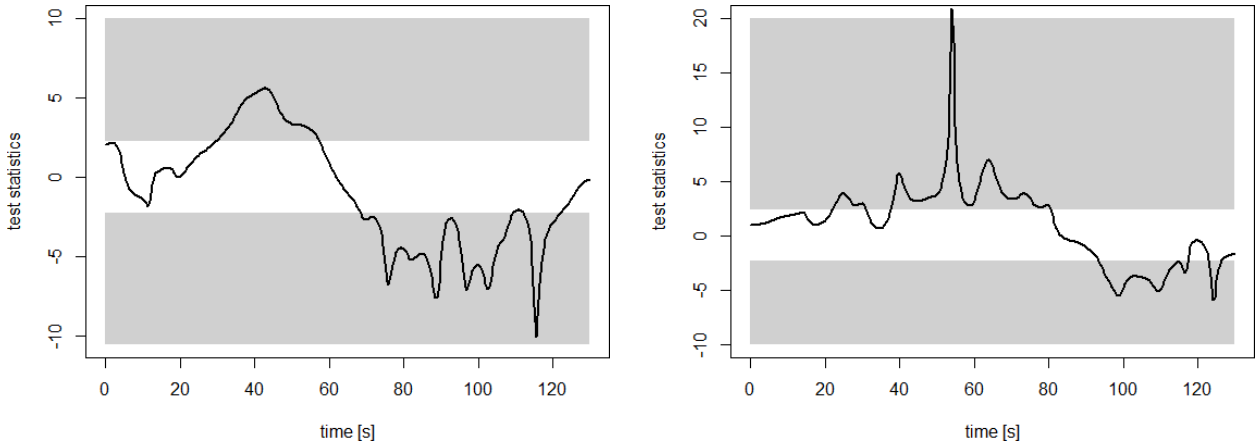
The graphical result of the positivity test of the derivative of the test substance A is shown in Figure 5(a). There is presented a critical region given by the quantiles of Student distribution (in grey) and the resulting test statistics in its functional form. We can see that the derivative is positive in the interval starting at 30 seconds. It indicates that the response is plausible after 30 seconds on average from the beginning of the measurement.

Let us get a closer look to the other two statements from the thesis. For the substance A: "In a measurement time interval from 30 to 102 seconds, the response of the instrument varies within a certain response interval, which is considered sufficient to demonstrate the presence of the test substance." and for the substance B: "The response of the the detector lies in the time interval from



(a) substance A (b) substance B

Fig. 4. Mean functions with their confidence bands



(a) substance A (b) substance B

Fig. 5. Test statistics (black line) and the critical regions (grey bands)

55 to 117 seconds." While being difficult to read from the graphs in Figure 5, the statements can be confirmed from the numerical values of the derivative test, see Table 1.

For the substance A, the derivative is different from zero mainly in the interval from 30 to 109 seconds (including the interval nearby the maximum), and for the substance B, from 38 to 117 seconds. The very short intervals at the beginning and at the end of the measurement can be omitted, because these are the places where the standard deviation is large, see Figure 4.

Thus, we confirm the statements from the thesis. First, the plausible interval starts at 30 seconds after beginning of the measurement of the substance A. Next, the whole interval, where the response is considered sufficient, ends at 109 seconds. The difference between the stated value (102 s) and calculated value (109 s) is negligible considering that the whole measurement takes approximately two minutes. Last, for the substance B, the plausible interval ranges from 38 to 117 seconds. Again,

Tab. 1. Intervals of positivity and negativity of the derivatives

substance	positive derivative	negative derivative
A	$(30, 57)$	$(69, 109), (113, 121)$
B	$(23, 31), (38, 80)$	$(94, 117), (124, 126)$

the difference between the stated value (55 s) and calculated value (38 s) is larger than in the case of the substance A, however, it also can be considered negligible with respect to the length of the measurement [1].

CONCLUSION

The dissertation thesis is the final work of postgraduate studies, which represent the highest level of university studies. From the point of view of the education system, it is the most important written document in the student's study effort. Therefore, it has to be of adequate quality; moreover, it has to bring new pieces of knowledge to the studied field. Last but not least, the students demonstrate their ability to process a selected research problem with an interdisciplinary approach. This can be significantly supported by the students' cooperation not only with their supervisors but also with other experts.

With the assistance of available scientific methods, in our case using the statistical approach, we presented the way in which, with help of a consultant (statistics expert), it is possible to move some dissertation theses to a higher scientific level and thus contribute to the fulfilment of the demanding conditions imposed on theses.

References

[1] Bland, J. M., Altman, D. G. Statistical methods for assessing agreement between two methods of clinical measurement. *International Journal of Nursing Studies* Vol. 47, No. 8, 2010, p. 931–936. DOI 10.1016/j.ijnurstu.2009.10.001

[2] Cleveland, W. S. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, Vol. 74, No. 368, 1979, p. 829–836.

[3] Hastie, T. J., Tibshirani, R. J. *Generalized additive models*. Boca Raton: Chapman and Hall, 1990.

[4] Koška, P. *Surface contamination control by using thermal desorption* (Dissertation thesis). Brno: University of Defence, 2015.

[5] Lacinová, V., Karpíšek, Z. Gradient and line estimates employed in surveys. *Mendel*, Vol. 2014, No. January, 2014, p. 407–414.

[6] Leflerová, G. *Functional data analysis* (Student research paper). Brno: University of Defence, 2021.

[7] Maturo, F., Hošková-Mayerová, Š. Analyzing research impact via functional data analysis: A powerful tool for scholars, insiders, and research organizations. In: *Proceedings of the 31st*

International Business Information Management Association Conference Innovation Management and Education Excellence through Vision 2020. IBIMA, 2018, p. 1832–1842.
[8] Primerano, P., Milazzo, M. F., Risitano, F. Sustainability improvement of chemical processes using amino ionic liquids as catalysts in water. In: *ICHEAP12: 12th International Conference On Chemical & Process Engineering*. Chemical Engineering Transactions, 2015, p. 2275–2280. DOI 10.3303/CET1543380
[9] R Core Team. *R: A language and environment for statistical computing*, Vienna: R Foundation for Statistical Computing, 2018.
[10] Ramsay, J. O., Silverman, B. W. *Functional data analysis.* New York: Springer, 2005.
[11] Tabaszewski, M., Szymański, G. M. Engine valve clearance diagnostics based on vibration signals and machine learning methods. *Eksploatacja i Niezawodnosc – Maintenance and Reliability*, Vol. 22, No. 2, 2020, p. 331–339. DOI 10.17531/ein.2020.2.16
[12] Vališ, D., Pokora, O., Koláček, J. System failure estimation based on field data and semi-parametric modeling. *Engineering Failure Analysis*, Vol. 101, 2019, p. 473–484. DOI 10.1016/j.engfailanal.2019.04.014
[13] Wang, J. L., Chiou, J. M., Müller, H. G. Functional data analysis. *Annual Review of Statistics and Its Application*, Vol. 3, No. 1, 2016, p. 257–295. DOI 10.1146/annurev-statistics-041715-033624
[14] Zhang, J.-T. *Analysis of variance for functional data.* Boca Raton: CRC Press, 2014.

Acknowledgement

The work presented in this paper has been supported by the University of Defence, project "DZRO FVL".

CONSTRUCTION ON AN INFINITE CYCLIC MONOID OF DIFFERENTIAL NEURONS

Jan Chvalina
Brno University of Technology, Faculty of Electrical Engineering and Communication,
Department of Mathematics
Technická 10, 616 00 Brno, Czech Republic
chvalina@feec.vutbr.cz

Michal Novák
Brno University of Technology, Faculty of Electrical Engineering and Communication,
Department of Mathematics
Technická 10, 616 00 Brno, Czech Republic
novakm@feec.vutbr.cz

Bedřich Smetana
University of Defence, Faculty of Military Leadership,
Department of Quantitative Methods
Kounicova 65, 662 10 Brno, Czech Republic
bedrich.smetana@unob.cz

Abstract: *When regarding the structure of the most commonly used artificial neural networks – multilayer perceptron ones – and when discussing the functionality of artificial neurons, one can use a certain analogy with relations between descriptions of differential equations of a certain type. In this contribution we analyze powers of differential neurons and present a construction of a countably infinite cyclic semigroup of powers of artificial differential neurons as the basis of a further investigation of some other structures.*

Keywords: Artificial differential neuron, cyclic semigroup, powers of differential neurons.

INTRODUCTION

Though semigroups are very simple structures (sets with one associative binary operation) the algebraic theory of semigroups belongs to classical algebraic structure theories with deep meaning and numerous applications – cf. [1, 4, 8, 10, 12, 16]. If a is any element of a semigroup $(S, \cdot)$, then the subsemigroup $\langle a \rangle$ of $(S, \cdot)$ is generated by a and consists of all the positive integral powers of a:

$$\langle a \rangle = \{a, a^2, a^3, \dots\}.$$

If $\langle a \rangle = S$, then $(S, \cdot)$ is called a cyclic semigroup. In a general case, we say that $\langle a \rangle$ is the cyclic subsemigroup of $(S, \cdot)$ generated by a. There are only two possibilities:

(1) No two powers of a are equal. Then evidently the element a has (countably) infinite order.

(2) There exist positive integers r and s with $r < s$ such that $a^r = a^s$. Then a has a finite order.

Recall e.g. in the theory of semigroup algebras one result of Amitsur [1]: if $(S, \cdot)$ is the infinite cyclic semigroup generated by x, then the algebra $\Phi[S^1]$ (where $S^1 = S \cup e$) of S^1 over a field Φ is the ring of polynomials $\Phi[x]$ in x over Φ. See also [4] p. 159. Numerous further interesting results can be found in the above mentioned classical monograph [4].

In this contribution we construct the semigroup of differential neurons of infinite order. Moreover, we construct the neutral element corresponding to the neuron $D^0Ne_p(\vec{w})$ in order to obtain a monoid, in fact. This structure is in [4] denoted by S^1. Consideration obtained in this contribution lead to constructions of cyclic hypergroups; for these see e.g. [5, 11, 15, 20, 21]

Let us begin with the fact that a neuron, called also artificial or formal neuron, is the basic stone of the mathematical model of any neural network. It is to be noted that its design and functionality are derived from observations of biological neurons which are basic building blocks of biological neural networks such as the brain, spinal cord and peripheral ganglia. In the case of artificial – formal neurons, the information comes into the body of the neuron via inputs that are weighted, i.e., each input can be individually multiplied by a weight. The body of an artificial neuron then sums the weighted inputs and bias and "processes" the sum with a transfer function – cf. [2, 3, 6, 9, 10, 13, 17, 18, 19].

At the end, an artificial neuron passes the processed information via outputs (output functions). One can say that artificial neural networks can be viewed as weighted directed graphs in which artificial neurons are nodes and directed edges with weights are connections between neuron inputs and neuron outputs. Recall that in the framework of artificial neural networks there are networks of simple neurons called perceptrons. The basic concept of a single perceptron was introduced by Rosenblatt in 1958. Perceptrons compute single outputs (the output function) from multiple real-valued inputs by forming a linear combination according to its input weights, and then possibly putting the output through some nonlinear activation function. Mathematically, this can be written as

$$y(t) = \varphi\left(\sum_{i=1}^{n} w_i(t)x_i(t) + b\right) = \varphi\left(\vec{w}^T(t)\vec{x}(t) + b\right),$$

where $\vec{w}(t) = (w_1(t), \ldots, w_n(t))$ denotes the vector of time dependent weight functions, $\vec{x}(t) = (x_1(t), \ldots, x_n(t))$ is the vector of time dependent (or time varying) input functions, b is the bias and φ is the activation function. The use of time varying functions as weights and inputs is a certain generalization of the classical concept of artificial neurons from the work of Warren McCulloch and Walter Pitts (1943); see also references mentioned above.

1 DIFFERENTIAL NEURONS AND THEIR OUTPUT FUNCTIONS

In what follows, we will consider a certain generalization of classical artificial neurons mentioned above such that inputs x_i and weights w_i will be functions of argument t belonging into a linearly ordered (tempus) set T with the least element 0. As the index set we use the set $\mathbb{C}(J)$ of all continuous functions defined on an open interval $J \subset \mathbb{R}$. So, denote by W the set of all non-negative functions $w : T \to \mathbb{R}$ forming a subsemiring of the ring of all real functions of one real variable $x : \mathbb{R} \to \mathbb{R}$. Denote by $Ne(\vec{w}_r) = Ne(w_{r1}, \ldots, w_{rn})$ for $r \in \mathbb{C}(J)$, $n \in \mathbb{N}$ the mapping

$$y_r(t) = \sum_{k=1}^{n} w_{r,k}(t)x_{r,k}(t) + b_r$$

which will be called the artificial neuron with the bias $b_r \in \mathbb{R}$. By $\mathbb{AN}(T)$ we denote the collection of all such artificial neurons.

Neurons are usually denoted by capital letters X, Y or X_i, Y_i, nevertheless we use also notation $Ne(\vec{w})$, where $\vec{w} = (w_1, \ldots, w_n)$ is the vector of weights.

We suppose – for the sake of simplicity – that transfer functions (activation functions) φ, σ(or f) are the same for all neurons from the collection $\mathbb{AN}(T)$ or that this function is the identity function $f(y) = y$.

Now, similarly as in the case of the collection of linear differential operators, we will construct a group and hypergroup of artificial neurons. Concerning the concept of a hypergroup, see e.g. [7, 9, 14, 15, 20].

Denote by δ_{ij} the so called Kronecker delta, $i, j \in \mathbb{N}$, i.e., $\delta_{ii} = \delta_{jj} = 1$ and $\delta_{ij} = 0$, whenever $i \neq j$.

Suppose $Ne(\vec{w}_r)$, $Ne(\vec{w}_s) \in \mathbb{AN}(T)$, $r, s \in \mathbb{C}(J)$, $\vec{w}_r = (w_{r1}, \ldots, w_{r,n})$, $\vec{w}_s = (w_{s1}, \ldots, w_{s,n})$, $n \in \mathbb{N}$. Let $m \in \mathbb{N}$, $1 \leq m \leq n$ be a such an integer that $w_{r,m} > 0$. We define

$$Ne(\vec{w}_r) \cdot_m Ne(\vec{w}_s) = Ne(\vec{w}_u),$$

where

$$\vec{w}_u = (w_{u,1}, \ldots, w_{u,n}) = (w_{u,1}(t), \ldots, w_{u,n}(t)),$$
$$w_{u,k}(t) = w_{r,m}(t)w_{s,k}(t) + (1 - \delta_{m,k})w_{r,k}(t), t \in T$$

and, of course, the neuron $Ne(\vec{w}_u)$ is defined as the mapping $y_u(t) = \sum_{k=1}^{n} w_k(t)x_k(t) + b_u$, $t \in T$, $b_u = b_r b_s$. Further for a pair $Ne(\vec{w}_r)$, $Ne(\vec{w}_s)$ of neurons from $\mathbb{AN}(T)$ we put

$$Ne(\vec{w}_r) \leq_m Ne(\vec{w}_s),\ w_r = (w_{r,1}(t), \ldots, w_{r,n}(t)),\ w_s = (w_{s,1}(t), \ldots, w_{s,n}(t))$$

if $w_{r,k}(t) \leq w_{s,k}(t)$, $k \in \mathbb{N}$, $k \neq m$ and $w_{r,m}(t) = w_{s,m}(t)$, $t \in T$ and with the same bias.

Remark 1. *A certain generalization of the formal (artificial) neuron can be obtained from linear differential operators of the n-th order. Recall the expression of formal neuron with inner potential* $y_{-in} = \sum_{k=1}^{n} w_k(t)x_k(t)$, *where* $\vec{x}(t) = (x_1(t), \ldots, x_n(t))$ *is the vector of inputs,* $\vec{w}(t) = (w_1(t), \ldots, w_n(t))$ *is the vector of weights. Using the bias* b *of the considered neuron and the transfer function* σ *we can expressed the output as* $y(t) = \sigma\left(\sum_{k=1}^{n} w_k(t)x_k(t) + b\right)$.

Now consider a tribal function $u : J \to \mathbb{R}$, *where* $J \subseteq \mathbb{R}$ *is an open interval; inputs are derived from the function* $u \in \mathbb{C}^n(J)$ *as follows:*
$x_1(t) = u(t), x_2 = \frac{du(t)}{dt}, \ldots, x_n(t) = \frac{d^{n-1}(t)}{dt^{n-1}}$, $n \in \mathbb{N}$. *Further the bias* $b = b_0\frac{d^n u(t)}{dt^n}$. *As weights we use continuous functions* $w_k : J \to \mathbb{R}$, $k = 1, \ldots, n-1$.

Then formula

$$y(t) = \sigma\left(\sum_{k=1}^{n} w_k(t)\frac{d^{k-1}u(t)}{dt^{k-1}} + b_0\frac{d^n u(t)}{dt^n}\right)$$

is a description of the action of the neuron Dn *which will be called a formal (artificial) differential neuron. This approach allows to use solution spaces of corresponding linear differential equations.*

2 PRODUCTS AND POWERS OF DIFFERENTIAL NEURONS

Suppose $\vec{w}(t) = (w_1(t), \ldots, w_n(t))$ are fixed vectors of continuous functions $w_k : \mathbb{R} \to \mathbb{R}$ and b_0 be the bias for any polynomial $p \in \mathbb{R}_s[t]$, $n \leq s$, $s \in \mathbb{N}_0$. We consider a differential neuron $DNe_p(\vec{w})$ by the action

$$y_1(t) = \sum_{k=1}^{n} w_{1,k}(t) \frac{d^{k-1}p(t)}{dt^{k-1}} + b_0 \frac{d^n p(t)}{dt^n}$$

with the identity activation function $\varphi(u) = u$. According to the formula, we can calculate the output function of the differential neuron $D^2Ne_p(\vec{w}) = DNe_p(\vec{w}) \cdot DNe_p(\vec{w})$.

The product of neurons $Ne(\vec{w}_r) \cdot Ne(\vec{w}_s) = Ne(\vec{w}_u)$; outputs of neurons

$$y_r(t) = \sum_{k=1}^{n} w_{r,k}(t)x_k(t) + b_r,\ y_s(t) = \sum_{k=1}^{n} w_{s,k}(t)x_k(t) + b_s.$$

The vector of weights of the neuron $Ne(\vec{w}_u)$ is of the form $\vec{w}_u(t) = (w_{u,1}, \ldots, w_{u,n})$, where

$$w_{u,k}(t) = w_{r,m}(t)w_{s,k}(t) + (1 - \delta_{m,k})w_{r,k}(t),\ t \in T \text{ and } 1 \leq m \leq n.$$

Then the neuron $Ne(\vec{w}_u)$ is defined as the function $y_u(t) = \sum\limits_{k=1}^{n} w_{u,k}(t)x_k(t) + b_r b_s,\ t \in T$.

In more detail:

$$\begin{aligned}
w_{u,1}(t) &= w_{r,m}(t)w_{s,1}(t) + w_{r,1}(t),\\
w_{u,2}(t) &= w_{r,m}(t)w_{s,2}(t) + w_{r,2}(t),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
w_{u,m}(t) &= w_{r,m}(t)w_{s,m}(t),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
w_{u,n}(t) &= w_{r,m}(t)w_{s,n}(t) + w_{r,n}(t).
\end{aligned}$$

Application of the above product onto the case of differential neurons: $DNe_p(\vec{w}_r),\ DNe_p(\vec{w}_s)$ with output functions $y_r(t) = \sum\limits_{k=1}^{n} w_{r,k}(t)\frac{d^{k-1}p(t)}{dt^{k-1}} + b_r\frac{d^n p(t)}{dt^n},\ y_s(t) = \sum\limits_{k=1}^{n} w_{s,k}(t)\frac{d^{k-1}p(t)}{dt^{k-1}} + b_s\frac{d^n p(t)}{dt^n}$, where $p \in \mathbb{R}_l[t]$, $n \leq l$. Denote $DNe_p(\vec{w}_u) = DNe_p(\vec{w}_r) \cdot DNe_p(\vec{w}_s)$. Then the output function of the neuron $DNe_p(\vec{w}_u)$ has the form

$$y_u(t) = \sum_{\substack{k=1\\k\neq m}}^{n} \left(w_{r,m}(t)w_{s,k}(t) + w_{r,k}(t)\right) \frac{d^{k-1}p(t)}{dt^{k-1}} + w_{r,m}(t)w_{s,m}(t)\frac{d^{m-1}p(t)}{dt^{m-1}} + \qquad (*)$$

$$+ b_r b_s \left(\frac{d^n p(t)}{dt^n}\right)^2.$$

Now, using the above formula we can express output functions of powers $D^2Ne_p(\vec{w}_r)$, $D^\alpha Ne_p(\vec{w}_r)$ (for $\alpha \in \mathbb{N}$) and $D^0Ne_p(\vec{w}_r)$ (the neutral element – unit) of the infinite cyclic group $\{D^\alpha Ne_p(\vec{w}_r);\ \alpha \in \mathbb{Z}\}$. The output function $y_u^{[2]}(t)$ of the differential neuron is of the form

$$y_u^{[2]}(t) = \sum_{\substack{k=1\\k\neq m}}^{n}(w_{r,m}(t)+1)w_{r,k}(t))\frac{d^{k-1}p(t)}{dt^{k-1}} + w_{r,m}^2(t)\frac{d^{m-1}p(t)}{dt^{m-1}} + b_r^2\left(\frac{d^np(t)}{dt^n}\right)^2 =$$

$$= (w_{r,m}(t)+1)\sum_{\substack{k=1\\k\neq m}}^{n} w_{r,k}(t))\frac{d^{k-1}p(t)}{dt^{k-1}} + w_{r,m}^2(t)\frac{d^{m-1}p(t)}{dt^{m-1}} + b_r^2\left(\frac{d^np(t)}{dt^n}\right)^2.$$

Theorem 1. *Consider a differential neuron $DNe_p(\vec{w})$ with the vector $\vec{w}(t) = (w_1(t), \dots, w_n(t))$ of time variable weights and the vector of inputs $\vec{x}(t) = (p(t), \frac{dp(t)}{dt}, \dots, \frac{d^np(t)}{dt^n})$ with polynomial $p \in \mathbb{R}_l[t]$, $n \leq l$, $t \in T$ and $1 \leq m \leq n$, $n \in \mathbb{N} = \{1, 2, \dots\}$. The output function $y(t)$ of the above mentioned neuron is of the form*

$$y(t) = \sum_{k=1}^{n} w_k(t)\frac{d^{k-1}p(t)}{dt^{k-1}} + b\frac{d^np(t)}{dt^n}$$

with the bias $b\frac{d^np(t)}{dt^n}$. Suppose $\alpha \in \mathbb{N}$, $2 \leq \alpha$. Then the output function of the differential neuron $D^\alpha Ne_p(\vec{w})$ has the form

$$y^{[\alpha]}(t) = \sum_{k=0}^{\alpha-1} w_m^k(t)\sum_{\substack{k=1\\k\neq m}}^{n} w_k(t)\frac{d^{k-1}p(t)}{dt^{k-1}} + w_m^{\alpha-1}(t)\frac{d^{m-1}p(t)}{dt^{m-1}} + \left(b\frac{d^np(t)}{dt^n}\right)^\alpha.$$

Proof. Consider a differential neuron $DNe_p(\vec{w})$ with the output function

$$y(t) = \sum_{k=1}^{n} w_k(t)\frac{d^{k-1}p(t)}{dt^{k-1}} + b\frac{d^np(t)}{dt^n}.$$

For $\alpha = 2$ we had obtained above the form of the output function

$$y^{[2]}(t) = (w_m(t)+1)\sum_{\substack{k=1\\k\neq m}}^{n} w_k(t)\frac{d^{k-1}p(t)}{dt^{k-1}} + w_m^2(t)\frac{d^{m-1}p(t)}{dt^{m-1}} + \left(b\frac{d^np(t)}{dt^n}\right)^2.$$

Suppose $\alpha \in \mathbb{N}$, $2 \leq \alpha$ and

$$y^{[\alpha]}(t) = \sum_{k=0}^{\alpha-1} w_m^k(t)\sum_{\substack{k=1\\k\neq m}}^{n} w_k(t)\frac{d^{k-1}p(t)}{dt^{k-1}} + w_m^{\alpha-1}(t)\frac{d^{m-1}p(t)}{dt^{m-1}} + \left(b\frac{d^np(t)}{dt^n}\right)^\alpha.$$

We will calculate the output function of the differential neuron $D^{\alpha+1}Ne_p(\vec{w})$. According to the formula $(*)$ we have $w_{s,k}(t) = \sum\limits_{k=0}^{\alpha-1} w_m^k(t)w_k(t)$ and thus

$$y^{[\alpha+1]}(t) = \sum_{\substack{k=1\\k\neq m}}^{n}\left(\left(w_m(t)\sum_{k=0}^{\alpha-1} w_m^k(t)\right)w_k(t) + w_k(t)\right)\frac{d^{k-1}p(t)}{dt^{k-1}} + w_m(t)w_m^{\alpha-1}(t)\frac{d^{m-1}p(t)}{dt^{m-1}} +$$

$$+b\frac{d^n p(t)}{dt^n}\left(b\frac{d^n p(t)}{dt^n}\right)^{\alpha}=\sum_{\substack{k=1\\k\neq m}}^{n} w_m(t)\left(\left(w_m^{\alpha-1}(t)+w_m^{\alpha-2}(t)+\cdots+w_m(t)+1\right)+1\right)w_k(t)\frac{d^{k-1}p(t)}{dt^{k-1}}+$$

$$+w_m^{\alpha}(t)\frac{d^{m-1}p(t)}{dt^{m-1}}+\left(b\frac{d^n p(t)}{dt^n}\right)^{\alpha+1}=$$

$$=\sum_{k=0}^{\alpha} w_m^k(t)\sum_{\substack{k=1\\k\neq m}}^{n} w_k(t)\frac{d^{k-1}p(t)}{dt^{k-1}}+w_m^{\alpha}(t)\frac{d^{m-1}p(t)}{dt^{m-1}}+\left(b\frac{d^n p(t)}{dt^n}\right)^{\alpha+1},$$

which is the output function for the power $D^{\alpha+1}Ne_p(\vec{w})$. The proof is complete. □

Example. Here we resent the output function of the power of the differential neuron $D^{\alpha}Ne_p(\vec{w})$ determined by values $n=5$, $m=4$ and $\alpha=3$. Moreover, consider the polynomial $p(t)=t^5+a_4t^4+a_3t^3+a_2t^2+a_1t+a_0$, a_k, $t\in\mathbb{R}$, i.e., $p(t)=\sum\limits_{k=0}^{5}a_kt^k$ with $a_5=1$.

We have $\vec{w}(t)=(w_1(t),\ldots,w_4(t),\ w_5(t))$, $t\in T$, which is the vector function of weights, thus the output function of the neuron $DNe_p(\vec{w})$ has the form

$$y(t)=\sum_{k=1}^{5} w_k(t)\frac{d^{k-1}p(t)}{dt^k}+b\frac{d^5p(t)}{dt^5}.$$

Since $\frac{dp(t)}{dt}=5t^4+4a_4t^3+3a_3t^2+2a_2t+a_1$, $\frac{d^2p(t)}{dt^2}=20t^3+12a_4t^2+6a_3t+2a_2$, $\frac{d^3p(t)}{dt^3}=60t^2+24a_4t+6a_3$, $\frac{d^4p(t)}{dt^4}=120t+24a_4$, $\frac{d^5p(t)}{dt^5}=120$, we obtain

$$y(t)=w_1(t)(t^5+\sum_{k=1}^{4}a_kt^k+a_0)+w_2(t)(5t^4+\sum_{k=1}^{3}ka_kt^{k-1})+w_3(t)(20t^3+\sum_{k=1}^{2}(k+1)ka_{k+1}t^{k-1})+$$

$$+6w_4(10t^2+4a_4t+a_3)+24w_5(t)(5t+a_4)+120b.$$

Then from Theorem 2 there follows that the output function of the neuron $D^3Ne_p(\vec{w})$ has the form

$$y^{[3]}(t)=\sum_{k=0}^{2} w_4^k(t)\sum_{\substack{k=1\\k\neq 4}}^{5} w_k(t)\frac{d^{k-1}p(t)}{dt^{k-1}}+w_4^2(t)\frac{d^3p(t)}{dt^3}+(b\frac{d^5p(t)}{dt^5})^3=$$

$$=(w_4^2(t)+w_4(t)+1)[w_1(t)(t^5+\sum_{k=1}^{4}a_kt^k+a_0)+w_2(t)(5t^4+\sum_{k=1}^{3}ka_kt^{k-1})+$$

$$+w_3(t)(20t^3+\sum_{k=1}^{2}(k+1)ka_{k+1}t^{k-1})+24w_5(t)(5t+a_4)]+6w_4(t)(10t^2+4a_4t+a_3)+(120b)^3,$$

where $(120b)^3=1728000b^3$.

3 IDENTITY ELEMENT AND CYCLIC SEMIGROUP OF DIFFERENTIAL NEURONS

In what follows, we will construct a differential neuron which is the neutral element with respect to the product operation "$\cdot$" defined for neurons $D^{\alpha}Ne_p(\vec{w})$, $\alpha \in \mathbb{N}$.

Consider a differential neuron $DNe_p(\vec{w})$, where $\vec{w}(t) = (w_1(t), \ldots, w_m(t), \ldots, w_n(t))$, with the output function

$$y(t) = \sum_{k=1}^{n} w_k(t) \frac{d^{k-1}p(t)}{dt^{k-1}} + b_0 \frac{d^n p(t)}{dt^n},$$

with the bias b_0 and $p \in \mathbb{R}_s[t]$, $n \leq s$, $s \in \mathbb{N}$. We denote by $N1(\vec{e})_m$ (instead of $D^0Ne_p(\vec{w})$) a neuron such that

$$N1(\vec{e})_m \cdot_m DNe_p(\vec{w}) = DNe_p(\vec{w}) \cdot_m N1(\vec{e})_m = DNe_p(\vec{w}).$$

Here we have $\vec{e}(t) = (e_1, e_2, \ldots, e_m, \ldots, e_n)$, where $e_k = 0$ for any $k \in \{1, 2, \ldots, n\}$, $k \neq m$, $e_m = 1$ and with the bias $b_1 = 1$. The output function corresponding to the neuron $N1(\vec{e})_m$ is of the form

$$y_1(t) = \frac{d^{m-1}p(t)}{dt^{m-1}} + 1$$

for any $p \in \mathbb{R}_s[t]$, $t \in T, n \leq s$.

Theorem 2. *Let $DNe_p(\vec{w})$ be a differential neuron, with the vector of weights $\vec{w}(t) = (w_1(t), \ldots, w_m(t), \ldots, w_n(t))$, the output function*

$$y(t) = \sum_{k=1}^{n} w_k(t) \frac{d^{k-1}p(t)}{dt^{k-1}} + b_0 \frac{d^n p(t)}{dt^n},$$

with the bias b_0 and $p \in \mathbb{R}_s[t]$, $n \leq s$, $s \in \mathbb{N}$.
Then

$$DNe_p(\vec{w}) \cdot_m N1(\vec{e})_m = DNe_p(\vec{w}) = N1(\vec{e})_m \cdot_m DNe_p(\vec{w}). \tag{1}$$

Proof. Consider differential neurons $DNe_p(\vec{w}_r)$, $DNe_p(\vec{w}_v)$ and denote

$$DNe_p(\vec{w}_u) = DNe_p(\vec{w}_r) \cdot_m DNe_p(\vec{w}_v).$$

Further, let $\vec{w}_v(t) = (w_{v,1}(t), \ldots, w_{v,m}(t), \ldots, w_{v,n}(t))$. According to the above defined product of neurons, the output function of the neuron $DNe_p(\vec{w}_u)$ is of the form

$$y_u(t) = \sum_{\substack{k=1 \\ k \neq m}}^{n} w_{u,k}(t) \frac{d^{k-1}p(t)}{dt^{k-1}} + w_{u,m}(t) \frac{d^{m-1}p(t)}{dt^{m-1}} + b_r b_v \left(\frac{d^n p(t)}{dt^n} \right)^2,$$

thus

$$y_u(t) = \sum_{\substack{k=1 \\ k \neq m}}^{n} \left(w_{r,m}(t) w_{v,k}(t) + w_{r,k}(t) \right) \frac{d^{k-1}p(t)}{dt^{k-1}} + w_{r,m}(t) w_{v,m}(t) \frac{d^{m-1}p(t)}{dt^{m-1}} + b_r b_v \left(\frac{d^n p(t)}{dt^n} \right)^2.$$

Now, consider the product of neurons $DNe_p(\vec{w})$ (with $\vec{w}(t) = (w_1(t), \ldots, w_m(t), \ldots, w_n(t))$ and the bias b) and $N1(\vec{e})_m$. Denote $DNe_p(\vec{w}_u) = DNe_p(\vec{w}) \cdot_m N1(\vec{e})_m$. Suppose $y(t)$ is the output function of $DNe_p(\vec{w})$. Then according to the above equalities the output function of the differential neuron $DNe_p(\vec{w}_u)$ has the form

$$y_u(t) = \sum_{\substack{k=1 \\ k\neq m}}^{n} \left(w_m(t) \cdot 0 + w_k(t)\right) \frac{d^{k-1}p(t)}{dt^{k-1}} + w_m(t)\frac{d^{m-1}p(t)}{dt^{m-1}} + b\frac{d^n p(t)}{dt^n} =$$

$$= \sum_{\substack{k=1 \\ k\neq m}}^{n} w_k(t)\frac{d^{k-1}p(t)}{dt^{k-1}} + w_m(t)\frac{d^{m-1}p(t)}{dt^{m-1}} + b\frac{d^n p(t)}{dt^n} =$$

$$= \sum_{k=1}^{n} w_k(t)\frac{d^{k-1}p(t)}{dt^{k-1}} + b\frac{d^n p(t)}{dt^n} = y(t),$$

hence $DNe_p(\vec{w}_u) = DNe_p(\vec{w})$.

Further, denote $DNe_p(\vec{w}_v) = DN1(\vec{e})_m \cdot_m DNe_p(\vec{w})$. With respect to the above consideration we have the output function y_v of the neuron $DNe_p(\vec{w}_v)$ in the form

$$y_v(t) = \sum_{\substack{k=1 \\ k\neq m}}^{n} \left(e_m w_k + e_k\right) \frac{d^{k-1}p(t)}{dt^{k-1}} + e_m w_m(t)\frac{d^{m-1}p(t)}{dt^{m-1}} + b\frac{d^n p(t)}{dt^n} =$$

$$= \sum_{\substack{k=1 \\ k\neq m}}^{n} \left(w_k(t) + 0\right) \frac{d^{k-1}p(t)}{dt^{k-1}} + w_m(t)\frac{d^{m-1}p(t)}{dt^{m-1}} + b\frac{d^n p(t)}{dt^n} =$$

$$= \sum_{k=1}^{n} w_k(t)\frac{d^{k-1}p(t)}{dt^{k-1}} + b\frac{d^n p(t)}{dt^n} = y(t),$$

for any $t \in T$, i.e., $DNe_p(\vec{w}_v) = DNe_p(\vec{w})$. Consequently, we obtain

$$DNe_p(\vec{w}) \cdot_m N1(\vec{e})_m = DNe_p(\vec{w}) = N1(\vec{e})_m \cdot_m DNe_p(\vec{w}).$$

□

Corollary. *Let $DNe_p(\vec{w})$ be a differential neuron, $m, n \in \mathbb{N}$, $1 \leq m \leq n$. Then for any positive integer $\alpha \in \mathbb{N}$ we have*

$$D^\alpha Ne_p(\vec{w}) \cdot_m N1(\vec{e})_m = D^\alpha Ne_p(\vec{w}) = N1(\vec{e})_m \cdot_m D^\alpha Ne_p(\vec{w}), \tag{2}$$

where $p \in \mathbb{R}_s[t]$, $n \leq s$, $s \in \mathbb{N}$, $t \in T$.

Proof. We use method of the mathematical induction.

I. Suppose that $\alpha = 1$. Then equalities (2) for $\alpha = 1$ are given by the above Theorem 2.

II. Suppose that $\alpha \in \mathbb{N},\ 1 \leq \alpha$ and equalities (2) hold. Then, with respect to Theorem 2,

$$D^{\alpha+1}Ne_p(\vec{w}) \cdot_m N1(\vec{e})_m = D^{\alpha}Ne_p(\vec{w}) \cdot_m DNe_p(\vec{w}) \cdot_m N1(\vec{e})_m =$$

$$= D^{\alpha}Ne_p(\vec{w}) \cdot_m DNe_p(\vec{w}) = D^{\alpha+1}Ne_p(\vec{w}) = DNe_p(\vec{w}) \cdot_m D^{\alpha}Ne_p(\vec{w}) =$$

$$= N1(\vec{e})_m \cdot_m DNe_p(\vec{w}) \cdot_m D^{\alpha}Ne_p(\vec{w}) = N1(\vec{e})_m \cdot_m D^{\alpha+1}Ne_p(\vec{w}).$$

□

As a result, we obtain that

$$(\mathbb{S}^1, \cdot_m) = (\{N1(\vec{e})_m\} \cup \{D^{\alpha}Ne_p(\vec{w});\ \alpha \in \mathbb{N}\},\ \cdot_m)$$

is a countable infinite cyclic semigroup with the identity (neutral element) $N1(\vec{e})_m$ – i.e., a monoid in fact – generated by the differential neuron $DNe_p(\vec{w})$.

CONCLUSION

As we have already mentioned in the Introduction, semigroups are significant and applicable algebraic structures in spite of their simplicity. In connection with the theory of neural networks and the theory of linear ordinary differential operators, we have constructed a semigroup with an identity element of differential neurons. This is the first step to construction of cyclic groups of the mentioned neurons and to constructions of cyclic hypergroups discussed e.g. in [5, 11, 15, 20, 21].

Author Contributions: Contributions of all authors of this paper are equal.

References

[1] Amitsur, A. S.: *Selected papers of S. A. Amitsur with commmentary*, Avinoam Mann, Amitai Regev, Louis Rowen, David .I. Saltman, and Lance W. Small (Editors), American Mathematical Society, Providence, RI, 2001.

[2] Behnke, S.: *Hierarchical Neural Networks for Image Interpretation*, Notes in Computer Science, Springer, Heidelberg, 2003.

[3] Bishop, C. M.: *Neural Networks for Pattern Recognition.* Oxford University Press, 1995.

[4] Clifford, A. H., Preston, G. B.: *The Algebraic Theory of Semigroups I.*, Amer. Math. Soc., Providence, 1977.

[5] De Salvo, M., Freni, D.: Sugli ipergruppi ciclici e completi. *Matematiche* (Catania) 1980, 35, 211–226.

[6] Hagan, M., Demuth, H., Beale, M.: *Neural Network Design*, PWS Publishing, Boston, MA, 1996.

[7] Hošková-Mayerová, Š., Chvalina, J.: Discrete transformation hypergroups and transformation hypergroups with phase tolerance space. *Discrete mathematics*, 2008, 4133–4143.

[8] Chvalina, J., Chvalinová, L.: Action of join spaces of continuous function on the underlying hypergroups of 2-dimensional linear spaces of functions. In *Aplimat 2009* Bratislava: STU Bratislava, 2009, 49–58.

[9] Chvalina, J., Smetana, B.: Groups and hypergroups of artificial neurons. In *17th CONFERENCE ON APPLIED MATHEMATICS APLIMAT 2018 PROCEEDINGS.* Bratislava: Slovak University of Technology in Bratislava, 2018, 232–243.

[10] Chvalina, J., Smetana, B.: Systems of fragments of artificial neural networks.. In *Proceedings, 19th Conference on Applied Mathematics Aplimat 2020.* Slovak University of Technology in Bratislava, 2020, 253–270.

[11] Karimian M., Davvaz, B.: On the γ-cyclic hypergroups. *Commun. Algebra* 2006, 34, 4579–4589.

[12] Ljapin, E. S.: *Semigroups*, American Mathematical Society, 1974.

[13] McCulloch, W., Pitts, W.: A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 1943, 5, 115–133.

[14] Novák, M., Cristea, I. Composition in *EL*–hyperstructures. *Hacet. J. Math. Stat.* 2019, 48, 45–58.

[15] Novák, M., Křehlík, Š., Cristea, I. Cyclicity in *EL*–hypergroups. *Symmetry* 2018, Vol.10, n.11, 1–13.

[16] Pondělíček, B.: *Algebraické struktury s binárními Operacemi* (Algebraic Structures with Binary Operations) SNTL, Praha 1977.

[17] Rosenblatt, F.: *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms.* Spartan, Washington DC, 1962.

[18] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. *J. Machine Learning Res.* 2014, 15, 1929–1958.

[19] Volná, E.: *Neuronové sítě 1. (Neural Networks 1.)* 2. ed., Ostravská univerzita, Ostrava, 2008.

[20] Vougiouklis, T.: Cyclicity in a special class of hypergroups. *Acta Univ. Carolinae Math. Phys.* 1981, 22, 3–6.

[21] Vougiouklis, T.: Isomorphisms on *P*–hypergroups and cyclicity. *Ars Combinatoria* 1990, 29A, 241–245.

ON THE SUMMABILITY OF NON-CONVERGENT SEQUENCES OF ELEMENTS OF BANACH SPACE

Alexander Maťašovský and Tomáš Visnyai
Faculty of Chemical and Food Technology, Slovak University of Technology in Bratislava
Radlinského 9, 812 37 Bratislava, Slovakia, alexander.matasovsky@stuba.sk,
tomas.visnyai@stuba.sk

Abstract: *The aim of this article is to show that by the regular matrix transformation could a divergent sequence of elements of Banach space become convergent.*

Keywords: Banach space, matrix transformation, convergence, summability.

INTRODUCTION

Let $(X, \|\cdot\|)$ be an arbitrary Banach space. The elements of Banach space we denote by $\alpha, \beta, \dots$, the zero element by Θ and the unit element by ϵ, where $\|\epsilon\| = 1$. The sequence $\alpha = (\alpha_n)$, $\alpha_n \in X$, $n = 1, 2, \dots$ converges to $\beta \in X$ if for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that for all $n > n_0$ implies $\|\alpha_n - \beta\| < \varepsilon$.

The sequence $\alpha = (\alpha_n)$, $\alpha_n \in X$, $n = 1, 2, \dots$ is called a Cauchy-sequence if for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that for all $i, j > n_0$ the inequality $\|\alpha_i - \alpha_j\| < \varepsilon$ holds.

It is well known, if $(X, \|\cdot\|)$ is a Banach space that every Cauchy sequence is convergent in X (see [3], [4] and [9]).

Let $A = (a_{mn})$ $(m, n = 1, 2, \dots)$ be an infinite matrix of real numbers. A sequence $\alpha = (\alpha_n)$, where $\alpha_n \in X$ for all $n = 1, 2, \dots$ is said to be A-*limitable* (limitable by a method (A)) to an element $a \in X$, if $\lim_{m\to\infty} \beta_m = a$, where $\beta_m = \sum_{n=1}^{\infty} a_{mn}\alpha_n$. If a sequence $\alpha = (\alpha_n)$ is A-limitable to the element a, we write $\text{A-lim}_{n\to\infty} \alpha_n = a$. The method (A) defined by the matrix A is said to be *regular* if $\lim_{n\to\infty} \alpha_n = a$ implies $\text{A-lim}_{n\to\infty} \alpha_n = a$ (see [9], [10]). If the method (A) is regular then the matrix A is called *regular transformation matrix* or in short *regular matrix*. For more details and examples let see [6] or [7].

The case of non-regular matrix transformation of sequences of elements of Banach space was investigated in [11]. There was proved that a matrix which transforms every bounded sequence into a convergent sequence of elements of Banach space i.e. *Schur matrix* can not be regular.

Recall, that the above-mentioned limitation method is a generalization of the notion of convergence of sequences. The latest results about convergence field of regular matrix transformation can be found in [12].

1 MAIN RESULT

For the sequences of elements of Banach space we have the following theorem.

Theorem 1. *Let $A = (a_{mn})$ be an infinite matrix of real numbers. The sequence*

$$\beta_m = \sum_{n=1}^{\infty} a_{mn}\alpha_n$$

converges to a for $m \to \infty$ and $\alpha_n \to a$ if and only if the following conditions hold:

a) there is a constant K such that $\sum_{n=1}^{\infty} |a_{mn}| \leq K$ for every $m = 1, 2, \ldots$,

b) for every $n = 1, 2, \ldots$, $\lim_{m\to\infty} a_{mn} = 0$,

c) $\lim_{m\to\infty} \sum_{n=1}^{\infty} a_{mn} = 1$.

Proof. See [10]. □

The infinite matrix $A = (a_{mn})$ of real numbers is regular if and only if it satisfies each condition of previous theorem (see [6]). In [5] was proved that the conditions of Theorem 1 are independent, i.e. by omitting any condition from Theorem 1 the matrix $A = (a_{mn})$ becomes non-regular.

In [2], the Steinhaus theorem is proved under the condition that there does not exist a regular matrix which limits all sequences of 0's and 1's. Now we will show an analogue of the Steinhaus theorem for sequences of elements of X. Let us define the set

$$\Omega = \{\alpha = (\alpha_k) \ : \ \alpha_k \in X, \ k = 1, 2, \ldots, \ \forall k = 1, 2, \ldots \ \|\alpha_k\| = 0 \ \vee \ \|\alpha_k\| = 1\},$$

which is the set of all sequences of elements of X with norm either zero or one (see [9]).

In the following theorem, we give a sufficient condition to the existence of at least one divergent sequence from Ω which is limitable by a matrix transformation. R. P. Agnew has given in [1] a similar simple sufficient condition that a regular matrix shall sum a divergent sequence of real numbers.

Theorem 2. *If the matrix $A = (a_{nk})$ is such that*

$$\sum_{k=1}^{\infty} |a_{nk}| < \infty, \quad n = 1, 2, \ldots, \tag{1}$$

$$\lim_{n\to\infty} \max_{k=1,2,\ldots} |a_{nk}| = 0, \tag{2}$$

then there is at least one divergent sequence, whose elements of space Ω, which is summable (limitable) by matrix $A = (a_{nk})$.

Proof. Let $A = (a_{nk})$ be a matrix satisfying (1) and (2). We show that a sequence of elements of space Ω, where $(X, \|\cdot\|)$ is a Banach space, will limitable by $A = (a_{nk})$ and its limit will the zero element Θ of space X.

Let $b_1, b_2, \dots$ be a sequence, of positive numbers, which converges to 0 so rapidly than nb_n (e.g. $b_n = \frac{1}{n^2}$). Let $c_1, c_2, \dots$ be a sequence, of positive numbers such that $\lim_{n\to\infty} c_n = 0$.

The condition (2) implies existence of an increasing sequence $n_1 < n_2 < \cdots$ of positive integers such that, for each $p = 1, 2, \dots$

$$|a_{nk}| \leq b_p, \quad n \geq n_p, \quad k = 1, 2, \dots . \tag{3}$$

Such a sequence (n_p) being fixed, the condition (1) implies that if $k_1 < k_2 < \cdots$ is a sequence of integers which becomes infinite sufficiently rapidly, then for each $p = 1, 2, \dots$

$$\sum_{k=k_p+1}^{\infty} |a_{nk}| \leq c_p, \quad n_p \leq n \leq n_{p+1}. \tag{4}$$

Let a sequence (k_p) be fixed such that (4) holds and $k_{p+1} > k_p + 1$ for each $p = 1, 2, \dots$. Let $\xi = (\xi_k) \in \Omega$ be the particular sequence of elements of Ω defined by the following way:

$$\xi_k = \begin{cases} \epsilon, & k = k_1, k_2, \dots, \\ \Theta, & \text{otherwise}, \end{cases}$$

where ϵ is the unit element of Ω and $\|\epsilon\| = 1$, Θ is the zero element of Ω and $\|\Theta\| = 0$.

The sequence $\xi = (\xi_k)$ is divergent according to the norm $\|\cdot\|$ in Ω. Moreover for the transformation of this sequence $\mu = (\mu_n)$ we have

$$\begin{aligned} \|\mu_n\| = \left\| \sum_{k=1}^{\infty} a_{nk}\xi_k \right\| &\leq \sum_{k=1}^{\infty} |a_{nk}| \, \|\xi_k\| \\ &= \sum_{j=1}^{\infty} \left|a_{nk_j}\right| \|\epsilon\| \\ &\leq \sum_{j=1}^{p} \left|a_{nk_j}\right| + \sum_{j=p+1}^{\infty} \left|a_{nk_j}\right| \\ &\leq \sum_{k=1}^{p} b_j + \sum_{k=k_p+1}^{\infty} |a_{nk}| \\ &< pb_p + c_p, \end{aligned}$$

where $p = 1, 2, \dots$ and $n_p \leq n < n_{p+1}$. Therefore since $pb_p \to 0$ and $c_p \to 0$ we have $\|\mu_n\| \to 0$. Therefore the sequence $\mu = (\mu_n)$ converges to Θ. Hence we found a divergent sequence of space Ω which is summable by matrix $A = (a_{nk})$. □

Example 1. Define the matrix $Z_{\frac{1}{2}} = (z_{mn})$ so that $z_{11} = \frac{1}{2}$, $z_{1n} = 0$ for $n = 2, 3, \ldots$, if $m = n$ then $z_{mn} = z_{mn+1} = \frac{1}{2}$ and finally $z_{mn} = 0$ otherwise. This matrix has the following form

$$Z_{\frac{1}{2}} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & \cdots & 0 & \cdots \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 & \cdots \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The matrix $Z_{\frac{1}{2}}$ satisfies each condition of Theorem 1 so it is regular but does not satisfy condition (2). Let $(L_1, \|\cdot\|)$ is the Banach space of Lebesgue integrable functions on the interval $[0, 1]$ with the norm $\|f\| = \int_0^1 |f(x)| \, \mathrm{d}x$. Let us define the sequence $f = (f_n)_{n=1}^{\infty}$, where $f_n \in L_1$ for all $n = 1, 2, \ldots$ in the following way:

$$f_{2k}(x) = \begin{cases} 1, & x \in \mathbb{Q} \cap [0, 1], \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q} \cap [0, 1], \end{cases}$$

$$f_{2k-1}(x) = \begin{cases} 0, & x \in \mathbb{Q} \cap [0, 1], \\ 1, & x \in \mathbb{R} \setminus \mathbb{Q} \cap [0, 1], \end{cases}$$

where $k = 1, 2, \ldots$. It is clear that the sequence $f = (f_n)_{n=1}^{\infty}$ does not converge with respect to the norm $\|\cdot\|$ in L_1. Create a sequence $\left(Z_{\frac{1}{2}} f_n\right)_{n=1}^{\infty}$. Then the transformed sequence we can express as

$$\begin{aligned} (g_m)_{m=1}^{\infty} &= \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & \cdots & 0 & \cdots \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 & \cdots \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ \vdots \end{pmatrix} \\ &= \left(\frac{1}{2} f_1, \frac{1}{2} \left(f_2 + f_3 \right), \ldots, \frac{1}{2} \left(f_{2k} + f_{2k+1} \right), \frac{1}{2} \left(f_{2k+1} + f_{2k+2} \right), \ldots \right) \\ &= \left(\frac{1}{2} f_1, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \ldots \right). \end{aligned}$$

The sequence $g = (g_m)_{m=1}^{\infty}$ converges according to the norm $\|\cdot\|$ in L_1 to the constant function $g(x) = \frac{1}{2}$ for all $x \in [0, 1]$.

The previous example shows, that Theorem 2 gives only sufficient conditions to the existence of at least one divergent sequence in Ω which is summable by matrix $A = (a_{nk})$.

The above mentioned matrix is not Schur since it is regular (see [6]), but it transforms bounded sequence to convergent (see [11]).

In [9] was proved a generalization of Steinhaus theorem for sequences of a Banach space and showed that the result of [4] and [8] can be generalized for a space of sequences of element of a Banach space. In the following theorem, we show a more generalization for an arbitrary norm.

Theorem 3. *For any regular matrix $A = (a_{nk})$ there exists a sequence in the set Ω, which is not limitable (summable) by matrix $A = (a_{nk})$.*

Proof. The matrix $A = (a_{nk})$ is regular, so it satisfies the conditions *a)–c)* of Theorem 1. On the base of condition *c)* we can choose such an index n_1 that

$$\sum_{k=1}^{\infty} a_{n_1 k} > \frac{3}{4}.$$

Similarly, by using condition *a)* we can choose an index m_1 such that

$$\sum_{k=m_1+1}^{\infty} |a_{n_1 k}| < \frac{1}{24}.$$

Let us define the sequence $\alpha = (\alpha_k) \in \Omega$ in the following way:

$$\alpha_k = \begin{cases} \epsilon, & 1 \le k \le m_1, \\ \Theta \text{ or } \epsilon, & k > m_1, \end{cases}$$

where ϵ is the unit element of Ω with $\|\epsilon\| = 1$ and Θ is the zero element of Ω with $\|\Theta\| = 0$.

Consider a series

$$\beta_{n_1} = \sum_{k=1}^{\infty} a_{n_1 k} \alpha_k,$$

then for its norm we get

$$\begin{aligned} \|\beta_{n_1}\| = \left\| \sum_{k=1}^{\infty} a_{n_1 k} \alpha_k \right\| &\ge \sum_{k=1}^{\infty} a_{n_1 k} \|\epsilon\| - \sum_{k=m_1+1}^{\infty} |a_{n_1 k}| \, \|\alpha_k\| \\ &\ge \sum_{k=1}^{\infty} a_{n_1 k} - 2 \sum_{k=m_1+1}^{\infty} |a_{n_1 k}| \\ &> \frac{3}{4} - 2 \cdot \frac{1}{24} = \frac{2}{3}. \end{aligned}$$

By using the condition *b)* we can choose such an index n_2 that $n_2 > n_1$ and $\sum_{k=1}^{m_1} |a_{n_2 k}| < \frac{1}{6}$. Then we can find $m_2 > m_1$ for which $\sum_{k=m_2+1}^{\infty} |a_{n_2 k}| < \frac{1}{6}$ holds, through the condition *a)*. Now, if we define the sequence $\alpha = (\alpha_k) \in \Omega$ in the following way

$$\alpha_k = \begin{cases} \epsilon, & 1 \le k \le m_1, \\ \Theta, & m_1 < k \le m_2, \\ \Theta \text{ or } \epsilon, & k > m_2, \end{cases}$$

then for the norm of β_{n_2} we get

$$\|\beta_{n_2}\| = \left\| \sum_{k=1}^{\infty} a_{n_2 k} \alpha_k \right\| \le \sum_{k=1}^{m_2} a_{n_2 k} \|\epsilon\| + \sum_{k=m_2+1}^{\infty} |a_{n_2 k}| \, \|\alpha_k\| < \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

By continuing the steps, we can choose such an index n_3 that $n_3 > n_2$, $\sum_{k=1}^{\infty} a_{n_3 k} > \frac{3}{4}$ and $\sum_{k=1}^{m_2} |a_{n_3 k}| < \frac{1}{24}$. Then we can find $m_3 > m_2$ for which $\sum_{k=m_3+1}^{\infty} |a_{n_3 k}| < \frac{1}{48}$ holds. Now, if we define the sequence $\alpha = (\alpha_k) \in \Omega$ as

$$\alpha_k = \begin{cases} \epsilon, & 1 \leq k \leq m_1, \\ \Theta, & m_1 < k \leq m_2, \\ \epsilon, & m_2 < k \leq m_3, \\ \Theta \text{ or } \epsilon, & k > m_3, \end{cases}$$

then for the norm of β_{n_3} we get

$$\begin{aligned} \|\beta_{n_3}\| = \left\| \sum_{k=1}^{\infty} a_{n_3 k} \alpha_k \right\| &\geq \sum_{k=1}^{\infty} a_{n_3 k} \|\epsilon\| - \sum_{k=m_1+1}^{m_2} |a_{n_3 k}| \, \|\alpha_k\| - 2 \sum_{k=m_3+1}^{\infty} |a_{n_3 k}| \, \|\alpha_k\| \\ &> \frac{3}{4} - \frac{1}{24} - 2 \cdot \frac{1}{48} = \frac{2}{3}. \end{aligned}$$

If we follow the steps above, we construct such a sequence $\alpha = (\alpha_k) \in \Omega$, that for the norm of transformed sequence we get

$$\left\|\beta_{n_{2l-1}}\right\| > \frac{2}{3} \qquad \text{and} \qquad \|\beta_{n_{2l}}\| < \frac{1}{3}, \qquad l = 1, 2, \ldots .$$

Therefore the sequence $\beta = (\beta_n)$ is not convergent according to the norm $\|\cdot\|$. Then the sequence $\alpha = (\alpha_n)$ is not summable by the regular matrix $A = (a_{nk})$. □

CONCLUSION

In this article, we have shown some results about the method of matrix limitation of divergent sequences of an arbitrary Banach space, which by using a regular matrix transformation become convergent with respect to the norm. Finally, we have created an example, where we have shown the Theorem 2 gives only sufficient conditions to the existence of at least one divergent sequence in Banach space, which is summable by regular matrix. The Steinhaus theorem was also proved in a Banach space i.e. for any regular matrix we can find a sequence of elements of Banach space which is not limitable by the matrix.

References

[1] AGNEW, R. P. *A simple sufficient condition that a method of summability be stronger than convergence*, Bull. Amer. Math. Soc., Vol. 52, 1946, p. 128-132.

[2] CONNOR, J. *A short proof of Steinhaus' theorem on summability*, Amer. Math. Monthly, Vol. 92, 1985, p. 420-421.

[3] KELLEY, J. L. - NAMIOKA, I. *Linear Topological Spaces.* New York: Springer, 1963.

[4] KOSTYRKO, P. Convergence fields of regular matrix transformations. *Tatra Mt. Math. Publ*, Vol. 28, No. 2, 2004, p. 153-157. ISSN 1338-9750.

[5] MAŤAŠOVSKÝ, A. On regularity conditions of matrix transformations. *Mathematics, Information Technologies and Applied Sciences 2019, post-conference proceedings of extended versions of selected papers*. Brno: University of Defence, 2019, p. 67-73. [Online]. [Cit. 2021-04-20]. Available at: http://mitav.unob.cz/data/Mitav2019.pdf. ISBN 978-80-7231-123-2.
[6] PETERSEN, G. M. *Regular matrix transformations.* London: McGraw-Hill, 1966.
[7] ŠALÁT, T. *Infinite series* [in Slovak]. Praha: Academia, 1974.
[8] ŠALÁT, T. On convergence fields of regular matrix transformations. *Czechoslovak Mathematical Journal* Vol. 26, No. 4, 1976, p. 613-627. ISSN 0011-4642.
[9] VISNYAI, T. Convergence fields of regular matrix transformations of sequences of elements of Banach spaces. *Miskolc Mathematical Notes*, Vol. 7, No. 1, 2006, p. 101-108. ISSN 1787-2413.
[10] VISNYAI, T. Limitation of Sequences of Banach Space through Infinite Matrix. *Mathematics, Information Technologies and Applied Sciences 2017, post-conference proceedings of extended versions of selected papers*. Brno: University of Defence, 2017, p. 256-262. [Online]. [Cit. 2021-04-20]. Available at: http://mitav.unob.cz/data/MITAV2017Proceedings.pdf. ISBN 978-80-7582-026-6.
[11] VISNYAI, T. Transformation of elements of the Banach space by Schur matrix. *Mathematics, Information Technologies and Applied Sciences 2019, post-conference proceedings of extended versions of selected papers*. Brno: University of Defence, 2019, p. 109-115. [Online]. [Cit. 2021-04-25]. Available at: http://mitav.unob.cz/data/Mitav2019.pdf. ISBN 978-80-7231-123-2.
[12] ZAKAWAT, U. S. - Dauda I. N. *On Convergence Fields of Some Regular Matrix Transformations*, World Journal of Innovative Research (WJIR), Vol. 8, No. 5, 2020, p. 06-08.

Acknowledgement

The work presented in this paper has been supported by the Scientific Grant Agency VEGA No. 1/0267/21.

DIFFERENTIAL NON – ANTAGONISTIC GAME WITH ADDITIONAL PAYMENT

Mokhonko E.Z.
Dorodnicyn Computing Center FRC CSC RAS
Vavilova, 40, 119333 Moscow, Russia, ezmokhon@mail.ru

Abstract: *Some special strategies are used. They permit to receive information about positions as sample data or in continuous way. The first player is able to pay an additional payment to partner. The payment influences on the character of the regime of the information reception about the game going. The influence is investigated.*

Keywords: differential non – antagonistic game, information reception, equilibrium, additional payment.

INTRODUCTION

As a rule, the most favourable regime of the information reception exists for a system of the dynamical processes control. Under this regime the system controls the processes in a best way and does not depreciate prematurely. So, it is clear how important and actual to investigate the optimum regimes of information receipt by means of the dynamic games.
Chernousko F.L., Melikjan A.A., Kononenko A.F., Mokhonko E.Z., [1] investigated how receive the same result using the sample data information instead of the continuous reception of information.
In this paper some differential game is considered. The first player is able to pay additional payment to the second player, if the second player does not deviate from the agreed trajectory. The equilibrium situation is constructed in r-strategies with a possibility to pay an additional payment. The payment changes the character of information receipt about trajectory. For example, it is getting possible to receive information about the equilibrium trajectory not countable times but the finite number times only.
The aim of the article is to clarify the character of changes of the information receipt about equilibrium trajectory under real or supposed changes of additional payment.

I. DESCRIPTION OF THE DIFFERENTIAL GAME

Let consider some differential non - antagonistic game of two players

$$x' = f(x,t,u,v),\ t_0 \le t \le T, \tag{1}$$

$$x(t_0) = x^0, \tag{2}$$

$$u \in P,\ v \in Q, \tag{3}$$

$$I_1(u,v) = g_1(x(T)), I_2(u,v) = g_2(x(T)) + U(x(T)). \tag{4}$$

Here x is n- dimensional vector of state, u and v are p- and q- dimensional vector - functions of control. Players 1 and 2 choose the meaning of the functions in order to maximize the appropriate cost functions $I_1(u,v) = g_1(x(T)), I_2(u,v) = g_2(x(T)) + U(x(T))$.

Player 1 uses the control u, player 2 uses the control v,

g_1 and g_2 are the continuous functions.

The sets P and Q are compacts in the appropriate vector spaces. The vector-function $f(x,t,u,v)$ is continuous function of all its arguments and satisfies restrictions which are imposed on it in [2]. The condition of existence the saddle point in the little game [3] is fulfilled. f, u, v have n, p, q -components respectively.

$U(x(T)) = \begin{cases} U_0 \geq 0, x(T) = x^{u0} \\ 0, x(T) \neq x^{u0} \end{cases}$. If $U_0 > 0$, then $U(x(T))$ is additional payment. The player 1 pays it to the second player if the trajectory of game is x^{u0} at the end of the game. If $U_0 = 0$, then the game without the additional payment is considered.

The set of permissible strategies of every player $\bar{U}, \bar{V}$ is the set of measurable for every argument positional $u(x,t), v(x,t)$ and program $u(t), v(t)$ controls. In addition to this some special strategies $\bar{u}, \bar{v}$ are permissible. They are called $r-$strategies [1, 4].

Let remind the definition of r-strategies.

Definition: *$r-$strategy $\bar{u}$ of player 1 is called the pair $\bar{u} = (r, u(.))$. It puts into correspondence to every point (x,t) the non - negative number $r \geq 0$.*

If $r > 0$, then the pair puts into correspondence to (x,t) the measurable function $u(\theta)$, $u(.) = \{u(\theta) = u(x,t;\theta) \in P \mid t \leq \theta < t + r(x,t)\}$.

If $r = 0$, then the pair puts into correspondence to (x,t) the control in the point (x,t): $u(.) = u(x,t)$.

We will use the concepts of Euler's broken line and motion [3].

In supplement and [1,4] there are the notions of the moments of the information receipt for the Euler's broken lines and motions, which are born by r – strategies. In the case $U_0 = 0$, the strategies $\bar{u}^0, \bar{v}^0$ were found in [1]. It forms the equilibrium situation. It gives birth to the equilibrium trajectory $x^0(t)$. The number of the moments of information receipt for the motion $x^0(t)$ is not more than the countable number.

Definition:

The pair of r-strategies $\bar{u}^0, \bar{v}^0$ forms the equilibrium situation in the game (1)-(4) if

1) the controls $\bar{u}^0, \bar{v}^0$ bear the unique solution of the problem (1)-(2) which is the unique motion, that is for all t the next correlation takes place

$X[x^0, t, \bar{u}^0, \bar{v}^0] = \{x^0(\tau), t \leq \tau \leq T\}$,

2) the equalities are fulfilled

$$I_1(\bar{u}^0, \bar{v}^0) = \max_{X[x^0(t), t, \bar{v}^0]} g_1(x(T)),\ t \in [t_0, T],$$

$$I_2(\bar{u}^0, \bar{v}^0) = \max_{X[x^0(t), t, \bar{u}^0]} g_2(x(T)), t \in [t_0, T].$$

Let consider the case $U_0 = 0$ that is the game without the additional payment in detail.

Let choose some piece - constant $u^0(t), v^0(t), t \in [t_0, T]$ and corresponding to them trajectory $x^0(t) = (x_1^0(t), x_2^0(t))$.

Let designate

$M_1 = \{x,t \mid g_1(x(T)) \le g_1(x^0(T))\}$, $M_2 = \{x,t \mid g_2(x(T)) \le g_2(x^0(T))\}$.

G_1 is the maximum v - stable bridge [3] to the set M_1, ∂G_1 is it's boundary.

G_2 is the maximum u - stable bridge [3] to the set M_2, ∂G_2 is it's boundary.

u^{ext} is the strategy which is extreme to the bridge G_2.

Let explain that the bridge G_u to the set M is the set G_u which has the next properties.

1. Bridge G_u contains the initial position $\{x_*, t_*\}$.
2. Bridge G_u finishes in the terminal set M.
3. The strategy $\hat{u}(x,t)$ exists which keeps every motion $x(t; x_*, t_*, \hat{u}(x,t))$ under any choice of the initial point $\{x_*, t_*\} \in G_u$ on the bridge till the meeting with M.

Notice that the second player cannot deviate from M using his controls v.

Extreme strategy $u^{ext}(x,t)$ and u - stability of the bridge G_u are described in [3].

Lemma 15.1, [3], states.

Let $u^{ext}(x,t)$ is the extreme strategy to the u-stable bridge G_u and let $(x_*, t_*) \in G_u$. Then for every motion $x(t) = x(t; x_*, t_*, u^{ext}(x,t))$ up to the meeting $(x(T), T) \in M$ inclusion is realized $\{x(t), t\} \in G_u (t_0 \le t < T)$.

Let $x^0(t) \in ((G_1 \setminus \partial G_1) \cap (G_2 \setminus \partial G_2)), t \in [t_0, T)$.

$X[x,t,u^0(\tau)]$ is the set of motions which appear due to the control function $u^0(\tau), \tau \in [t,T]$ and begin from the point x,t. $x(\tau; x,t,u^0(\tau))$ is the motion beginning from point x,t and appearing due to control $u^0(\tau)$. Control of the second player is not interesting for us.

$\bar{T}(x,t) = \{\bar{t} \mid \bar{t} \in [t,T],$

$\exists x(\tau; x,t,u^0(\tau)) \in X[x,t,u^0(\tau)],$

$\exists t_2 : x(\tau; x,t,u^0(\tau)) \notin G_2 \forall \tau \in (\bar{t}, t_2), t_2 \in (\bar{t}, T]\}$

$\omega_0(x,t) = \inf_{\bar{t} \in \bar{T}(x.t)} \bar{t}$, if $(x,t) \in G_2$. $\omega_0(x,t) = t$, if $(x,t) \notin G_2$,

$\omega_c(x,t) = \omega_0(x,t) - t$.

$\omega_c(x,t)$ is the minimal interval of time which is necessary for the second player in order to reach the boundary ∂G_2 from point (x,t).

The r - strategies $\bar{u}^0, \bar{v}^0$ form the equilibrium situation,

Where $\bar{u}^0 = \begin{cases} r(x,t) = \omega_c(x,t), (x,t) \in G_2; \\ r(x,t) = 0, (x,t) \notin G_2; \\ \{u^0(\tau) \mid t \le \tau < t + r\}, \\ r(x,t) > 0, (x,t) \in G_2; \\ u^0(t), (x,t) \in G_2, r = 0; \\ u^{ext}(x,t), (x,t) \notin G_2 \end{cases}$.

$\bar{v}^0$ is defined similarly.

The pair $\bar{u}^0, \bar{v}^0$ gives birth to the equilibrium trajectory $x^0(t)$. The number of the moments of information receipt for the motion $x^0(t)$ is not more than the countable number.

2. DIFFERENTIAL GAME WITH ADDITIONAL PAYMENT

Let consider the case $U_0\left(x^0(T)\right) > 0$, $x^{u0} = x^0(T)$, that is the game with the additional payment. Let $M_2^{U_0} = \left\{x,t \mid g_2\left(x(T)\right) \le g_2\left(x^0(T)\right) + U_0\right\}$.

$G_2(U_0)$ is the maximum u- stable bridge to the set $M_2^{U_0}$, $\partial G_2(U_0)$ is it's boundary.

$\bar{T}(x,t;U_0) = \{\bar{t} \mid \bar{t} \in [t,T],$

$\exists x\left(\tau; x,t, u^0(\tau)\right) \in X\left[x,t,u^0(\tau)\right],$

$\exists t_2 : x\left(\tau; x,t,u^0(\tau)\right) \notin G_2(U_0) \forall \tau \in (\bar{t}, t_2), t_2 \in (\bar{t}, T]\}$

$\omega_0^{U0}(x,t) = \inf\limits_{\bar{t} \in \bar{T}(x,t;U_0)} \bar{t}$ if $(x,t) \in G_2(U_0)$, $\bar{T}(x,t;U_0) \ne \varnothing$.

$\omega_0^{U0}(x,t) = T$, if $(x,t) \in G_2(U_0)$, $\bar{T}(x,t;U_0) = \varnothing$.

$\omega_0^{U0}(x,t) = t$ if $(x,t) \notin G_2(U_0)$. $\omega_c^{U0}(x,t) = \omega_0^{U0}(x,t) - t$.

$\bar{u}^0(U_0) = \begin{cases} r(x,t) = \omega_c^{U0}(x,t), \\ (x,t) \in G_2(U_0); \\ r(x,t) = 0, (x,t) \notin G_2(U_0); \\ \{u^0(\tau) \mid t \le \tau < t + r\}, \\ r(x,t) > 0, (x,t) \in G_2(U_0); \\ u^0(t), (x,t) \in G_2(U_0), r = 0; \\ u^{ext}(x,t), (x,t) \notin G_2(U_0) \end{cases}$.

The next theorem is proved as it was done in [1].

Theorem1. The pair $\bar{u}^0(U_0), \bar{v}^0$ forms the equilibrium situation and gives birth to the equilibrium trajectory $x^0(t)$.

The number of the moments of information receipt for the motion $x^0(t)$ is finite number.

Let $U_{01}\left(x^0(T)\right) > U_{02}\left(x^0(T)\right) > 0$. The pair $\overline{u}^0(U_{01}), \overline{v}^0$ forms the equilibrium situation and gives birth to the equilibrium trajectory $x^0(t)$ as well as the pair $\overline{u}^0(U_{02}), \overline{v}^0$.

Theorem2. Let the pair $\overline{u}^0\left(U_{01}\right), \overline{v}^0$ gives birth to the motion $x^0(t)$. The amount of the moments of the information receipt by the player 1 about the motion $x^0(t)$ is not more than the amount of the moments of information receipt for the motion $x^0(t)$ which is born by the pair $\overline{u}^0(U_{02}), \overline{v}^0$.

The theorems 1 and 2 are able to be used for the investigation of the game, in which disturbance is able to change the value of addition payment.

Example.

$x'_1 = v$, $x'_2 = u$, $0 \le t \le 1$, $0 \le u \le 1$, $0 \le v \le 1$, $x_1(0) = 0$, $x_2(0) = 0$,

$I_1(u,v) = x_1(1) - \frac{1}{2}x_2(1)$, $I_2(u,v) = x_2(1) - \frac{1}{2}x_1(1) + U\left(x(T)\right)$.

The first player is player u, the second player is the player v.

At first let consider the case $U_0 = 0$ that is the game without additional payment.

The maximum u - stable bridge G_2 to the set $\left\{x, T \mid g_2(x) \le g_2\left(x^0(1)\right)\right\}$ is:

$$G_2 = \left\{(x, t \in A) \mid \min_{u(x,t)} \max_{X[x,t,u(x,t)]} g_2\left(x(1)\right) \le g_2\left(x^0(1)\right)\right\}, \text{ that is}$$

$$G_2 = \left\{(x, t \in A) \mid x_2(t) - x_1(t)\frac{1}{2} \le g_2\left(x^0(1)\right)\right\}.$$

Similarly

$$G_1 = \left\{(x, t \in A) \mid x_1(t) - x_2(t)\frac{1}{2} \le g_1\left(x^0(1)\right)\right\}, \text{ where}$$

$$A = \left\{(x,t) \mid (0 \le t \le 1) \wedge \left(0 \le x_1 \le 1\right) \wedge \left(0 \le x_2 \le 1\right)\right\}.$$

Let $x_1^0(t) = t$, $x_2^0(t) = t$, $g_1\left(x^0(1)\right) = \frac{1}{2}$, $g_2\left(x^0(1)\right) = \frac{1}{2}$.

$$\min_{u(x,t)} \max_{X\left[x^0(t),t,u(x,t)\right]} g_2\left(x(1)\right) = t - \frac{1}{2}t = \frac{1}{2}t < \frac{1}{2}, t \in [0,1),$$

$$\min_{v(x,t)} \max_{X\left[x^0(t),t,v(x,t)\right]} g_1\left(x(1)\right) = t - \frac{1}{2}t = \frac{1}{2}t < \frac{1}{2}, t \in [0,1),$$

That is $x^0(t) \in \left(G_2 \setminus \partial G_2\right) \wedge \left(G_1 \setminus \partial G_1\right)$ when $t \in [0,1)$. That is why the position strategies $u^0(x,t)$, $v^0(x,t)$ form the situation of equilibrium,

$$u^0(x,t) = \begin{cases} 1, x_1(t) \ge 2x_2^0(t) - 1 = 2t - 1 \\ 0, x_1(t) < 2x_2^0(t) - 1 = 2t - 1 \end{cases},$$

$$v^0(x,t)=\begin{cases}1, x_1^0(t)\le\frac{1}{2}x_2(t)+\frac{1}{2}\\ 0, x_1^0(t)>\frac{1}{2}x_2(t)+\frac{1}{2}\end{cases}$$, that is $$v^0(x,t)=\begin{cases}1, t\le\frac{1}{2}x_2(t)+\frac{1}{2}\\ 0, t>\frac{1}{2}x_2(t)+\frac{1}{2}\end{cases}$$.

Equilibrium strategies are

$$\overline{u}^0=\begin{cases}r(x,t)=\frac{1}{2}(x+1)-t,(x,t)\in G_2;\\ r(x,t)=0,(x,t)\notin G_2;\\ \quad\{u^0(\tau)=1\,|\,t\le\tau<t+r\},\\ \quad r(x,t)>0,(x,t)\in G_2;\\ u^0(t)=1,(x,t)\in G_2, r=0;\\ u^{ext}(x,t)=0,(x,t)\notin G_2\end{cases}$$

$\overline{v}^0$ is constructed similarly.

The moments of reception information by the first player about $x^0(t)$ are: $t_n=1-(0.5)^n, n=0,1,\dots, \lim\limits_{n\to\infty}t_n=1, \lim\limits_{n\to\infty}(t_n-t_{n-1})=\lim\limits_{n\to\infty}(0.5)^n=0$.

Let consider the case $U_0>0$. The first player pays it to the second player when $x_1^0(1)=1, x_2^0(1)=1$.

$$G_2(U_0)=\left\{(x,t\in A)\,|\,x_2(t)-x_1(t)\frac{1}{2}\le g_2\left(x^0(1)\right)+U_0\right\},$$

or $G_2(U_0)=\{(x,t\in A)\,|\,x_1(t)\ge 2t-1-2U_0\}$.

$$\overline{u}^0(U_0)=\begin{cases}r(x,t)=\frac{1}{2}(x+1+2U_0)-t,\\ (x,t)\in G_2(U_0);\\ r(x,t)=0,(x,t)\notin G_2(U_0);\\ \{u^0(\tau)=1\,|\,t\le\tau<t+r\},\\ r(x,t)>0,(x,t)\in G_2(U_0);\\ u^0(t)=1,(x,t)\in G_2(U_0), r=0;\\ u^{ext}(x,t)=0,(x,t)\notin G_2(U_0)\end{cases}$$

The moments of the information receipt for $x^0(t)$ are $t_n=(1+2U_0)\left(1-(0.5)^n\right)\ n=0,1,\dots,m(U_0)$. If $U_0=\frac{1}{2046}$, then $t_{10}=1$, $m(U_0)=m\left(\frac{1}{2046}\right)=10$. If $U_0=\frac{1}{126}>\frac{1}{2046}$, then $t_6=1, m(U_0)=m\left(\frac{1}{126}\right)=6$.

Consideration of the example is finished.

In conclusion we'll try to explain the results of the theorem 2 in the simplest way. So it is necessarily to do some designations and note some properties of the bridges.

$\rho\left(x^0(t),\partial G_2\right)\left(\rho\left(x^0(t),\partial G_2\left(U_{0i}\right)\right),i=1,2\right)$ is the distance between $x^0(t)$ and the boundary $\partial G_2\left(\partial G_2\left(U_{0i}\right),i=1,2\right)$ of the bridge $G_2\left(G_2\left(U_{0i}\right),i=1,2\right)$. It depends on t.

$U_{01}\left(x^0(T)\right)>U_{02}\left(x^0(T)\right)>0$. In common case $M_2\subset M_2^{U_{20}}\subset M_2^{U_{10}}$. According to the properties of bridges $G_2\subset G_2\left(U_{02}\right)\subset G_2\left(U_{01}\right)$. So

$\rho\left(x^0(t),\partial G_2\right)<\rho\left(x^0(t),\partial G_2\left(U_{02}\right)\right)<\rho\left(x^0(t),\partial G_2\left(U_{01}\right)\right)$.

The inequalities are used in conclusion.

CONCLUSION

In the paper it is shown that even the least additional payment arranges the finite quantity of the information reception about equilibrium trajectory $x^0(t)$. If additional payment increases then the amount of information reception decreases.

The frequency of reception of information about whether the second player deviated from $x^0(t)$ and reaches the boundary of the bridge depends on the distance between $x^0(t)$ and the boundary. In common case $\rho\left(x^0(t),\partial G_2\right)\underset{t\to T}{\longrightarrow}0$. So, the amount of the moments of the reception of information for $x^0(t)$ is countable.

$\rho\left(x^0(t),\partial G_2\left(U_{01}\right)\right)>\rho\left(x^0(t),\partial G_2\left(U_{02}\right)\right)>0, t\in\left[t_0,T\right]$. So, the amount of the moments of the reception of information for $x^0(t)$ is finite. If the distance is longer than the frequency is smaller.

It is known that in the game without additional payment [1, 2] $x^0(t)$ is born by equilibrium positional strategies. They demand the continuous reception of information.

SUPPLEMENT

In the supplement the definitions of the Euler's broken line $x_\Delta(t)=x_\Delta\left(t;x^*,t_*,\bar{u},v(.)\right)$ and motion $x(t)=x\left(t;x^*,t_*,\bar{u}\right)$, which are born by r-strategy $\bar{u}$ from position $\left(x^*,t_*\right)$, are adduced. Definitions of the moments of reception information for the Euler's broken line and motion which are born by r – strategy, are done.

Let the initial position $\left(x^*,t_*\right)$ is done and $r-$ strategy $\bar{u}$ is chosen. Let cover $\left[t_*,T\right]$ by the system of semi - intervals $\tau_i\le t<\tau_{i+1}, i\in I=\{0,1,...,n\}$, $\tau_0=t_*,\tau_n=T$. Let $v(t)\in Q,\left(t\ge t_*\right)$ is the measurable according to Lebesgue realization of the second player's control $v(.)=\left\{v(t)\,|\,t_*\le t\le T\right\}$. Then the Euler's broken line $x_\Delta(t)=x_\Delta\left(t;x^*,t_*,\bar{u},v(.)\right)$ is called absolutely continuous function which fits to the condition $x_\Delta\left(t_*\right)=x^*$ and which is the solution of the differential equation

$$x'_\Delta(t)=f\left(x_\Delta(t),t,u\left(x_\Delta\left(\tau_i\right),\left(\tau_i\right)\right),v(t)\right) \qquad (5)$$

under $\tau_i\le t<\tau_{i+1}$, if $i\in I_1$ and

$$x'_{\Delta}(t) = f\left(x_{\Delta}(t), t, u\left(x_{\Delta}\left(\tau_{j_0}\right), \tau_{j_0}; \tau_i\right), v(t)\right) \qquad (6)$$

under $\tau_i \le t < \tau_{i+1}$, if $i \in I_2$.

Here $u\left(x_{\Delta}\left(\tau_{j_0}\right), \tau_{j_0}; \tau_i\right)$ is meaning, which takes the function $u\left(x_{\Delta}\left(\tau_{j_0}\right), \tau_{j_0}; t\right)$ under $t = \tau_i$ (look the definition of $r-$ strategy).

The set of indices I_2 is defined in the following way.

If $i \in I_2$, then $1) i \ne 0$; $2) \exists j_0 < i$ such, that j_0 is not belongs to I_2 and $\tau_{j_0} + r\left(x_{\Delta}\left(\tau_{j_0}\right), \tau_{j_0}\right) > \tau_i$ (for every $i \in I_2$ such j_0 is unique).

The set of indices I_1 is $I_1 = I \setminus I_2$.

Let make clear this definition. For example, let $\tau_1 < \tau_0 + r\left(x^*, \tau_0\right) < \tau_2$. The function $u\left(x^*, t_*; t\right)$ is put into correspondence to the point $\left(x^*, t_*\right) = \left(x^*, \tau_0\right)$. Then the Euler's broken line fits to the condition $x_{\Delta}\left(t_*\right) = x^*$ and is the solution of the equation (5) under $\tau_0 \le t < \tau_1$ and the equation (6) under $\tau_1 \le t < \tau_2$, where $u\left(x_{\Delta}\left(\tau_{j_0}\right), \tau_{j_0}; \tau_i\right) = u\left(x^*, t_*; \tau_1\right)$.

If $\tau_1 \ge \tau_0 + r\left(x^*, \tau_0\right)$, then the Euler's broken line is the solution of the equation (5) under $\tau_0 \le t < \tau_1$, $\tau_1 \le t < \tau_2$. Further, the Euler's broken line fits to the equation (6) under $t \in \left[\tau_2, \tau_3\right)$, if $\tau_1 + r\left(x_{\Delta}(\tau_1), \tau_1\right) > \tau_2$. It fits to the equation (5) on the contrary case.

Definition:

The moment of time τ_j, $j \in I_1$, $j \ne 0$, is called the moment of reception of information for the Euler's broken line.

The motion $x(t) = x\left(t; x^*, t_*, \bar{u}\right)$, which is born by the strategy $\bar{u}$ from the position $\left(x^*, t_*\right)$, is called the every absolutely continuous function $x(t)$, for which the sequence of the Euler's broken line $x_{\Delta}^{(k)}\left(t; x^k, t^k, \bar{u}, v^k(.)\right)$ is found. The sequence evenly converges to $x(t)$ on $t_* \le t \le T$ under condition $\lim\limits_{k\to\infty} \sup\limits_i \left(\tau_{i+1}^k - \tau_i^k\right) = 0$.

Let remind that the distance between the trajectories $x_{\Delta}^k(t)$ and $x(t)$ is estimated by the equality

$$\left\|x_{\Delta}^{(k)}(t) - x(t)\right\| = \max\left\{\left\|x_{\Delta}^{(k)}(t) - x(t)\right\|_{C\left[t^k, 1\right]}, \left|t^k - t_*\right|\right\}, \text{ where}$$

$$\left\|x_{\Delta}^{(k)}(t) - x(t)\right\|_{C\left[t^k, 1\right]} = \max_{t^k \le t \le 1} \left\|x_{\Delta}^{(k)}(t) - x(t)\right\|.$$

The existence of motions which are born by r-strategy is proved in the same way as the existence of motions which are born by positional strategy is proved.

Definition:

Let the sequence of the Euler's broken lines $\chi_{\Delta}^{k}(t)$ converges to the motion $x(t)$. Let for the sequence of the Euler's broken lines the sequence of the moments of the reception of information exists, which converges to some moment t. Then the moment t is called the moment of reception of information for the motion $x(t)$ or the moment of reception of information simply.

By another words, if $\tau_i^k, i \in I_1^k, i \neq 0, \lim_{k\to\infty} \tau_i^k = t$, then t is the moment of reception of information for the motion $x(t)$. Here $I_1^k = I^k \setminus I_2^k$ is the subset of indices, which corresponds to the Euler's broken line $\chi_{\Delta}^{(k)}$.

References

[1] Kononenko A.F., Mokhonko E.Z,Processes of information reception in non-antagonistic differential game, *Reports on applied mathematics.* Moscow: CC AS USSR,1982,20pp.

[2] Kononenko A.F.,Structure of optimum strategies in dynamical control systems, *Journal of computer mathematics and mathematical physics.* 1980, vol. 20, №5, p. 1105-1116.

[3] Krasovsky N.N., Subbotin A.N. *Positional differential games,*Moscow: Nauka,1974, 456 pp.

[4] Mokhonko E.Z., Dynamics of information processes in nonantagonistic games, *Dissertation ...DSc. of physical and mathematical sciences: 05.13.09.* Moscow: CC RAS, 1998, 350 pp.

Implementation of the concept of active education in the field of technical subjects

Soňa Pavlíková[1], Michal Kuba[2], Dagmar Faktorová[3], Peter Fabo[4]
[1] FCPT STU, Radlinskeho 9, 812 37 Bratislava, sona.pavlikova@stuba.sk
[2,4] Fakulta špeciálnej techniky, TnUAD v Trenčíne, Ku kyselke 469, Trenčín
[3]Katedra merania a aplikovanej elektrotechniky, Univerzitná 8215/1, Žilina

Abstract: *The subject of the paper is the presentation of the concept of an active document to support the education of technical subjects. The concept is based on extending the capabilities of the Jupyter Notebook platform with the use of external software, such as simulators, graphic editors and component databases for the presentation and solution of problems. The paper outlines examples of possible use in the process of education these subjects: the basics of electronics, theoretical electrical engineering and circuit design.*

Keywords: jupyter notebook, simulations, python, education, electronic circuits

Introduction

The current state of multimedia resources utilization in classical education approach is focused mainly on the passive form of presentations in text form, supplemented by graphic elements, graphs, animations or videos. In the case of distance learning, this form is usually supplemented by a presentation of the lecturer in real time with a suitable form of mutual communication. The active form of the student's work is usually practice and solving tasks related to the material, which is supplemented by consultation with the teacher.

This classic method of education has limited technological possibilities. For this reason, it does not fully support creative and critical thinking, active research and experimentation with the presented topic. This has limited possibilities for the student to modify the presented problems, their starting points, parameters and the consequences of their changes. With the growing need to increase the quality of education with the support of self-study as well as distance learning, it is necessary to extend and re-evaluate the existing view of a simple presentation form of study materials. Advances in the development of computer tools enables the implementation of new approaches to education with the possibility of active student involvement in the subject matter, generally included under the active learning paradigm. A number of platforms are currently available for teacher interaction with students, such as Moodle, Chamilo, edX, Edmondo and others. These are primarily intended for communication in the form of video, chat and sharing of study materials.

For active study with the possibility of interaction with the study material, it is necessary to choose the form of active documents for that materials. In the field of creation, presentation and distribution of such documents, especially in the field of support for education subjects related to information technologies, the Jupyter Notebook platform [1] and projects derived from it for the creation of electronic publications [4] are widely used. The subject of this paper is the presentation of the

concept of extending this platform to the education of technical subjects - theoretical electrical engineering, design and simulation of electronic circuits, signal processing, which require the support of external software.

From the teacher point of view who is working in the field of teaching the electrical engineering, electronics and related subjects, which requires an active approach of the student to the presented topic, the concept of interactive documents should meet the following requirements:

- In the field of elementary teaching, visualization and practical verification of basic theoretical knowledge:
  - connection the mathematical description with the properties of the circuit,
  - the possibility of active work with the topic by its modification,
  - modification and addition the parameters (working with basic electrical circuits, response of basic RC elements to different types of signals in the time and frequency domain, etc.).
- In the field of electrical circuits analysis:
  - practical verification of results the theoretical investigation of the problem (loop currents, node voltages, admittance matrix, oriented graphs),
  - practical verification of results the analysis and design of the circuit with specific parameters and properties,
  - the possibility of investigation the properties in the domain not covered by the analysis (influence of real properties of components, extension by parasitic elements, etc.).
- In the field of synthesis and advanced circuit design:
  - connection the design methods with the knowledge obtained during the programming courses,
  - practical use of numerical algorithms in technical practice (circuits optimization for given parameters, analysis of component values variations for required properties, minimization of power losses, etc.).
- In the field of interface design, measurement and experiment control:
  - simulation the designed circuits properties using specific models of components (stability, noise properties, dynamic response, power parameters, etc.),
  - the possibility of realization the virtual experiments by simulation using mathematically synthesized signals as well as with signals obtained by measurements from the real world,
  - the possibility of proposed solutions optimization before their practical verification.

  For students the use of this concept can be extremely attractive for analysis, signal processing and design of circuit solutions in areas where access to real ”sources” of signals is technologically limited or impossible (biology, biophysics, medicine, nuclear technology, high voltage systems, transport, etc.)

In this contribution the possibility of extending the capabilities of Jupyter-Notebook platform by using a classic simulator of electronic circuits as an active part of study materials are discussed. The use of this concept on examples from the education of electronics is presented. Of course, this concept can be extended and implemented in the process of education in other areas of technical sciences, which require specific support for software that is not included in the Jupyter-Notebook infrastructure.

1 Jupyter Notebook

1.1 Properties

Over time, the Jupyter Notebook [1] platform has evolved from an interactive interpreter of the Python programming language into a large ecosystem for creating interactive platform-independent documents that use a web browser for presentation.

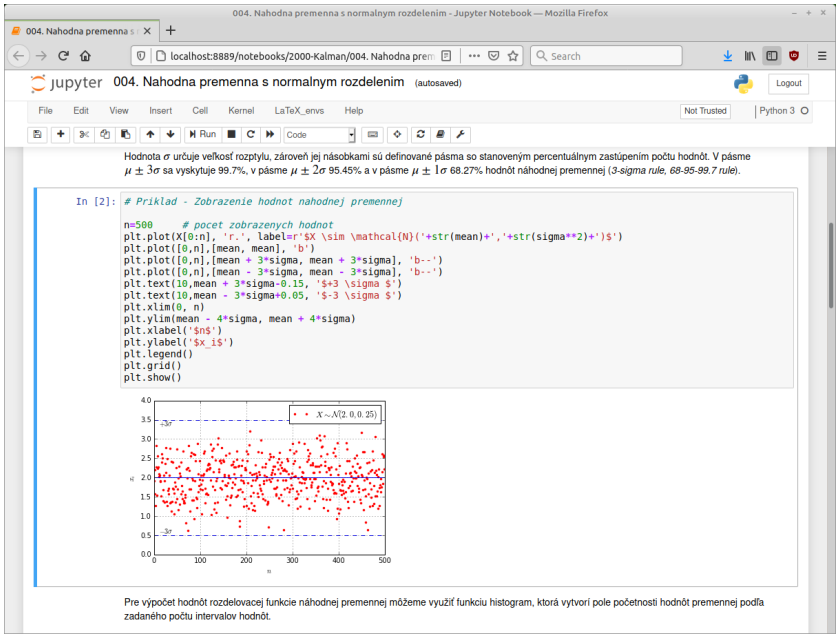


Figure 1: Jupyter Notebook in web browser

The possibilities of using this platform for teaching from the point of view of the teacher as well as the student are described in detail together with the case studies in [2]. The primary areas for the Notebook application are that, which overlap with information technology. Jupyter Notebook technology supports all the attributes of a standard web page, support for embedding executable codes in dozens of programming languages, rendering mathematical formulas and generating graphs. It is an ideal tool for creating textbooks and lecture notes in the area of programming, database systems, numerical mathematics, visualization and data processing. A number of examples are given e.g. in [3]. From a practical point of view the following points are important:

- The whole ecosystem is open-source, without restrictions and license fees, with the possibility of extensions and modifications.

- Creating and working with notebooks is platform independent in a web browser environment. The creation uses a simple markdown language with support for creating mathematical expressions using standard LaTeX syntax.
- Several notebooks distribution models are supported:
  - for a large number (> 100) of users via the *JupyterHub* server application [5],
  - for a smaller number (< 100) of users using the *Littlest JupyterHub* [6],
  - or individual use via a local multiplatform *jupyter* server emulation application,
  - for distribution of notebooks as passive web pages using the *nbviewer* [7] application or for conversion notebooks to an electronic book using *JupyterBook* [4] application with the possibility of exporting them to the classic paper form of the book or lecture notes.

1.2 Jupyter-Notebook expansion possibilities

The standard use of the *Jupyter Notebook* application does not require any intervention to the standard configuration of the student's computer. In the simplest case, just using a standard web browser is enough. A more complicated situation occurs when it is necessary to use a software within the notebook that is not a standard part of its ecosystem. The classic SPICE simulator was used in the presented concept. The possibilities of using computational programs for FDTD simulations of electromagnetic fields (MEEP [8]), interactive solution of problems using finite elements (gmsh, ElemerFEM [9]) including the use of access to remote systems were also tested.

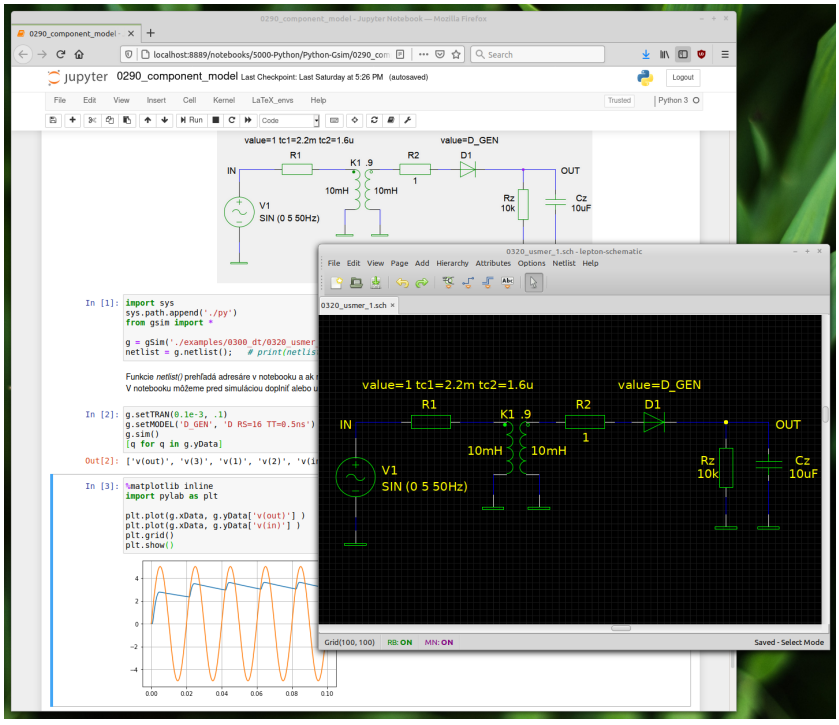


Figure 2: *Jupyter Notebook* with external *lepton-schematic* editor with circuit diagram

A typical configuration of a notebook page in the presented concept for solving electrical problems consists of the following parts:

- Theoretical part with a description of the problem, technically using the possibilities of notebooks (texts, pictures, mathematical relations, scripts for calculations, drawing graphs, tabulation of relations ...)
- Description of the circuit that is the subject of the problem. The input data containing the circuit description are entered into the *ngspice* [11] in a text file form in the SPICE format [10], which can be quite extensive and confusing for more complicated circuits. From a pedagogical point of view, it is more appropriate to use the classical graphical representation of the circuit in the form of its circuit diagram. Since there is currently no suitable graphical editor implemented directly in the notebook environment, in cases where there is the need to modify the circuit diagram, an external editor is used, which is called from the notebook environment (*lepton-schematic* [12]), Fig. 2.
- Simulation and presentation of results. In the presented concept of an interactive notebook with a demonstration of the properties of the electronic circuit, an external simulation engine ngspice was used and the results are processed and displayed in graphical form with the help of standard Python libraries (*pandas, matplotlib*).

The notebook extension infrastructure is shown in the following figure. Communication with the notebook extension is performed by means of the *gSim* interface, the description of which is the subject of the following section. The *ngspice* simulator itself is a standard program run in batch mode.

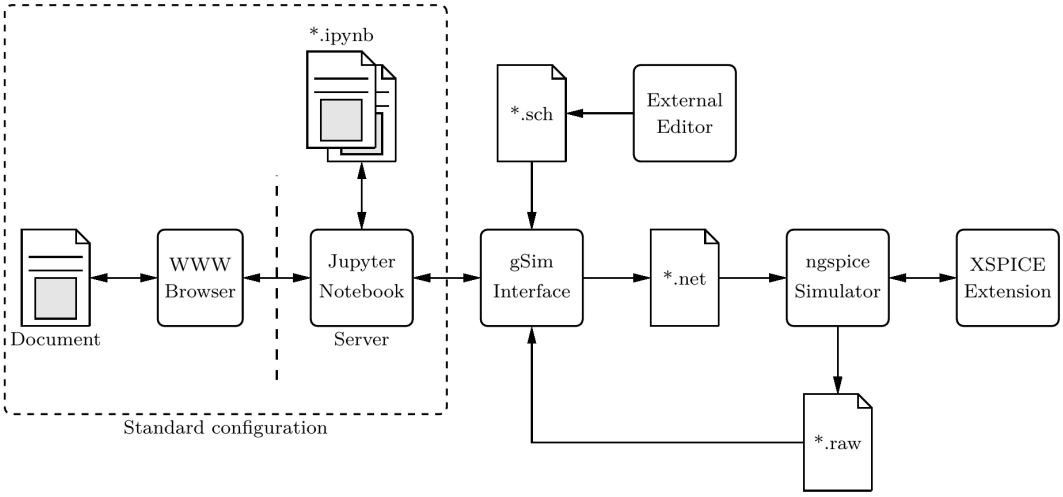


Figure 3: The extended *Jupyter Notebook* configuration

Before the actual simulation process it is necessary to connect the description of the circuit with internal databases and to add the models of the components from the local database. In the case of the block circuit simulation, e.g. differential equations using *XSPICE* components (part of *ngspice*), it is necessary to generate a specific model for each component with the current parameters specified in the circuit diagram. The result of the simulation is a binary or text file with voltages and currents values for the selected circuit nodes. These results need to be further converted into a data structures that can be visualized backwards in the notebook. A modified version of the gSim scripts was used to simplify the whole process of data preparation, simulation control and conversion of

results into data structures [13]. All these activities are implemented by software without the need for interaction with the user.

1.3 gSim library structure

The *gSim* scripts library is created in the *Python* programming language and forms a communication interface between the Notebook environment and external programs, the circuit editor, the netlist generator and simulator itself. It contains a set of simple methods for input data processing, simulation control and components parameter changing directly from the notebook environment.

1.3.1 Processing and editing of input data

The input data loading from the circuit diagram graphical editor is done by the `gSim` class constructor. At the same time a data consistency check is performed during loading (unconnected component pins, duplicate or unspecified component references, etc.). The circuit netlist conversion according to the SPICE format is done by the `netlist()` method.

```
from gsim import *                    # library import
g = gSim('./examples/test_1.sch')     # read and check schematics
nst = g.netlist()                     # create netlist
print(nst)                            # print netlist
```

It is appropriate to demonstrate the creation of a netlist by entering it manually for pedagogical reasons in order to understand the principle of the simulation process and possible circuit consistency checking.

```
netlist='* Komentar     \n' +\      # netlist as test string
'R1 1 2 1k              \n' +\      # in SPICE format
'C1 2 0 1uF             \n' +\
'V1 1 0 dc 0 ac 1       \n' +\
'.AC DEC 100 10Hz 1MEG  \n' +\
'.END                   \n'

fp = open('test_2.net', 'w')       # write to file
fp.write(netlist)
fp.close()

g = gSim()                         # create simulator object
g.netlist('test_2.net')            # read netlist from file
```

1.3.2 Simulation control

The created netlist can be edited and modified, the components can be added and updated, the parameter values can be changed and the simulation parameters can be set without the need to

interfere with the original circuit graphical input. Some of the methods for circuit modification and simulation control are listed below:

- `setDC()` setting DC simulation parameters
- `setAC()` setting AC simulation parameters
- `setTRAN()` setting TRANSIENT simulation parameters
- `setPAR()` setting the value of the circuit parameter
- `setOPT()` setting or changing a simulation parameter
- `setMODEL()` component model declaration
- `setCOMP()` adding the SPICE component to the circuit
- `getNET()` return the current form of the netlist
- `sim()` run simulation

The type and parameters of the simulation can be entered directly in the graphics editor or by a command before starting the simulation. The simulation result is written by the simulator to a binary or text file. Then it is transferred to a dictionary containing as a key a variable name and its associated data. The independent variable (voltage, time, frequency) is stored in the variable *xData*, the dependent variables (values of currents, voltages in the circuit nodes) are in the *yData* dictionary, while the name of the variable is the key of the dictionary.

```
g = gSim('test_3.sch')
g.netlist()
g.setTRAN(1e-5, 3, uic='UIC' )
g.sim()
[q for q in g.yData]     # list of output values
```

The last script command displays a list of available variables from the circuit simulation, voltages, and currents in the circuit nodes as dictionary keys. Depending on the type of simulation, the output is real or complex values.

```
['v(2)', 'v(3)', 'v(1)', 'v(out1)', 'i(v1)']
```

The result of the simulation can be further processed using standard Python libraries for data processing and visualization (*numpy*, *scipy*, *pandas*, *maptplotlib* ...).

```
import pandas as pd
import matplotlib.pylab as plt

d = pd.DataFrame(g.yData)          # create pandas data object
d.index = g.xData                  # from simulation result

plt.plot(d.index, d.xData['v(1)']) # create graph
plt.show()
```

2 Examples of use

The following section presents examples taken from selected areas of electronics and supplemented by comments with emphasis on the presentation of the properties of the described concept.

2.1 Basic electronics

Problem: *Based on the (previous) theoretical analysis of the bandpass RC filter (according to the following figure), display the transmission characteristics of the circuit for the values of the parameter R in the range of* 5 *to* $12\,k\Omega$*. Calculate the circuit properties in case the component values can be within a tolerance band of 5%. Compare the results with the previous solution for ideal component values.*

This example presents the use of parametric simulation of electronic circuits. The component values are given by the R, C, parameters which are substituted by specific values before the simulation itself.

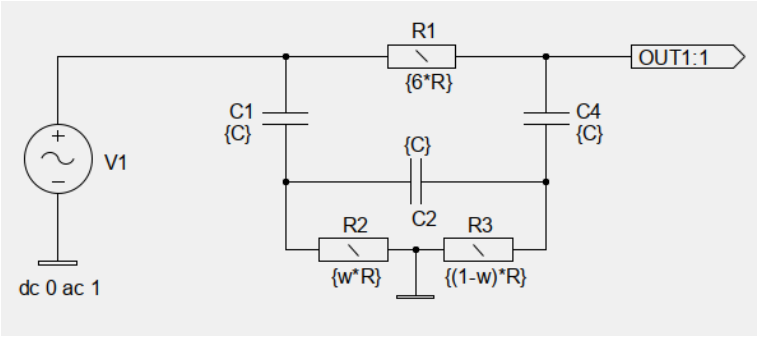


Figure 4: Parametric simulation of electronic circuits

The preparation of the simulation consists of loading a graphical representation of the bandpass filter created by the graphical editor and subsequent conversion to a text description of the circuit.

```
from gsim import *                      # library import
g = gSim('./examples/notch.sch')        # read and convert schematics
nst = g.netlist()                       # create netlist
print(nst)                              # print netlist
```

The script output is a text description of the circuit, a netlist in SPICE format, generated by a script from a circuit diagram. It is possible to compare the description of the circuit with its graphical representation.

```
R3 0 1 {(1-w)*R}
R2 3 0 {w*R}
C2 3 1 {C}
C1 2 3 {C}
V1 2 0 dc 0 ac 1
R1 2 OUT1 {6*R}
```

```
C4 OUT1 1 {C}
.end
```

The components parameter values can be set using the `setPAR()` method. The values can be entered in numeric or text form in compliance with SPICE format. Also the type of simulation and frequency range can be defined parametrically. Parameter values and the simulation process can be modified programmatically. he results of the AC simulation are values in a complex form. The circuit frequency response is displayed using standard *matplotlib* library.

```
g.setPAR('C', '0.1uF')    # text form of the parameter
g.setPAR('w', 0.055)      # numerical form of the parameter
                          # typa and range of simulation
g.setAC(10, 1000, number=10000, stype='DEC' )
for r in linspace(5e3, 12e3, 6):
    g.setPAR('R', r)      # parameter setting
    g.sim()
    plt.semilogy(g.xData, np.abs(g.yData['v(out1)']), label=str(r/1e3)+'k' )
...                       # creating graph, grid, legend ...
plt.show()
```

The output of the script are values from which a graph is generated with a set of magnitude frequency response characteristics of the circuit for a given range of parameters. The student can adjust the graph to a standard form with a decibel scale.

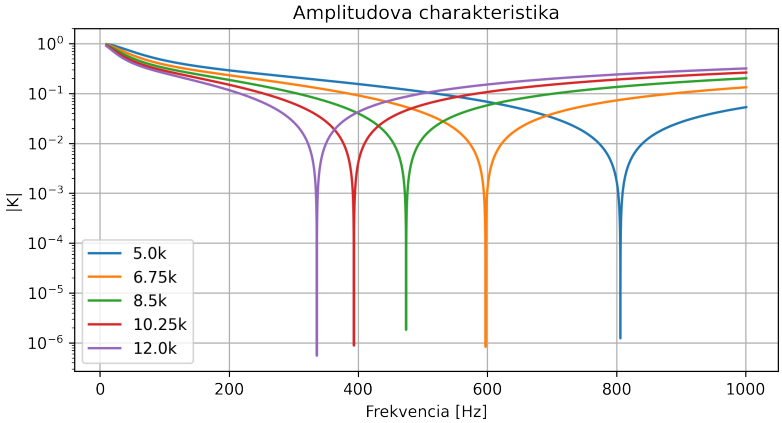


Figure 5: Results of parametric simulation from the first part of the problem

To solve the second part of the problem, it is necessary to introduce parameters for individual components and display the tolerance bands of characteristics by varying the values in the specified range.

2.2 Theoretical electronics

Problem: *Create and verify a model of a system of differential equations describing a harmonic oscillator. Create and verify the system model with operational amplifiers without using the multiplication operation.*

When presenting the fundamental properties of electronic circuits, it is appropriate to avoid from technical details and use the general terms to describe the problem.

A typical example is a harmonic oscillator which is described by a second-order differential equation:

$$\frac{\mathrm{d}^2 s}{\mathrm{d}t^2} + \omega^2\, s = 0$$

Using substitutions

$$y_1 = s \qquad \frac{\mathrm{d}y_2}{\mathrm{d}t} = \omega\, y_1$$

arrange the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}y_1}{\mathrm{d}t} = -\omega^2\, y_1 \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}y_1}{\mathrm{d}t} = -\omega\frac{\mathrm{d}}{\mathrm{d}t}y_2$$

and we obtain a system of two first-order differential equations

$$\frac{\mathrm{d}y_2}{\mathrm{d}t} = \omega\, y_1 \qquad \frac{\mathrm{d}y_1}{\mathrm{d}t} = -\omega\, y_2$$

for simulation in integral form

$$y_2(t) = \omega \int y_1(t)dt \qquad y_1(t) = -\omega \int y_2(t)dt$$

The system of integral equations can be displayed directly using blocks in the simulation diagram using functional blocks from the *XSPICE* extension which is part of the *ngspice* simulator. In circuit diagram with *XSPICE* blocks the work with them is as with standard electronic components and it is also possible to combine them freely.

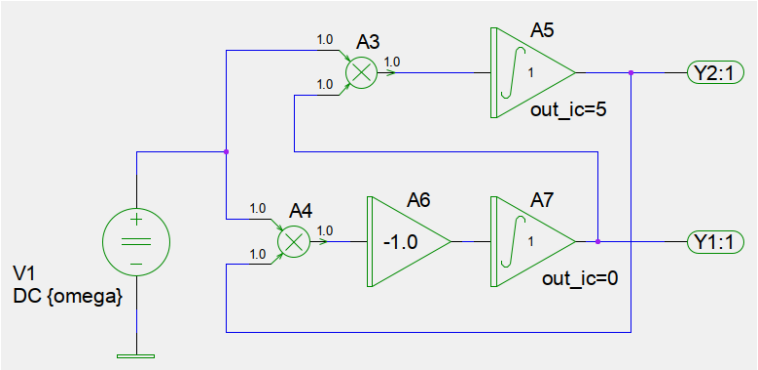


Figure 6: Simulation of a system of differential equations

In block diagram the voltage source determines the generated frequency. The initial condition of the integrator $A5$ determines the amplitude of the oscillations. The parameter ω sets the oscillator frequency, for the $\omega = 2\pi$ value the oscillator frequency will be 1Hz. The script contains the

setting of the simulation type (TRANSIENT) using the initial conditions (UIC) and the setting of the parameter determining the frequency.

```
g = gSim('./examples/generator.sch')
n = g.netlist()
g.setTRAN(1e-5, 3, uic='UIC' )
g.setPAR('omega', 2*pi)
g.sim()

plt.plot(g.xData, g.yData['v(y2)'], label=str('Y2'))
plt.plot(g.xData, g.yData['v(y1)'], label=str('Y1'))
plt.legend()
plt.grid()
plt.show()
```

The simulation output are waveforms on the circuit $Y1$ and $Y2$ ports with mutual phase shift which results from the theoretical description of the problem.

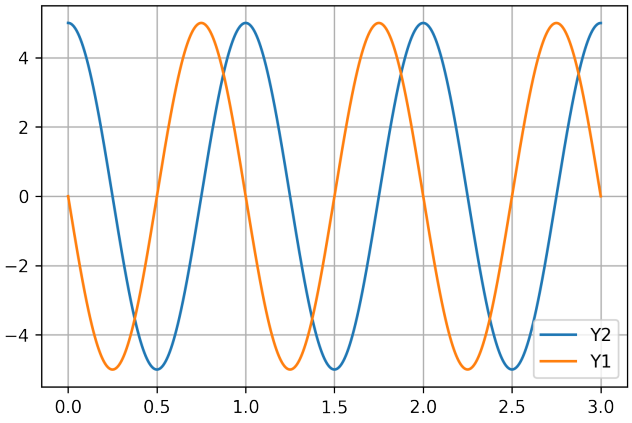


Figure 7: Time response at the output of signal generator

One of the possible solutions of the second part of the problem is the circuit according to the following figure:

2.3 Electronic circuit design

Problem: *Design the protection circuits of the ST32L432 microcontroller port which counts the number of mechanical contact closures. When verifying the circuit properties use the voltage waveform obtained by measuring on a real mechanical contact. For simulation purposes refer to the manufacturer's technical documentation for the microcontroller port circuit diagram.*

Unlike of simulators with non-public source code or commercial products, *ngspice* allows to create own components for *XSPICE* extensions and incorporate them into the simulation. For the following example, with a demonstration of the design of the microntroller port protection circuit, a

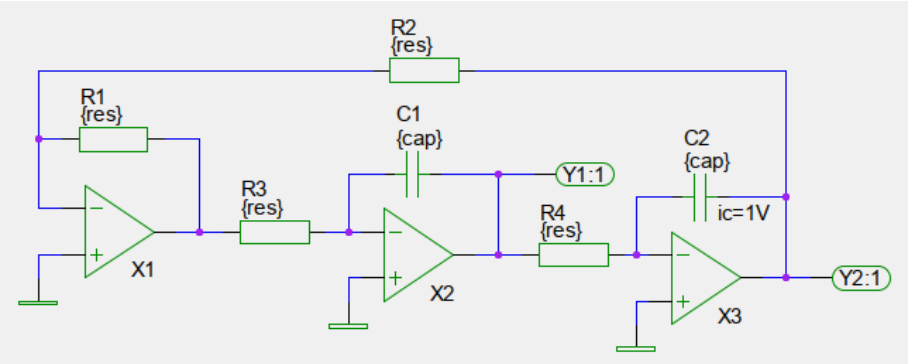


Figure 8: Simulation of a system of differential equations using electronic components

component was programmed that allows the simulation of the time response of the real switching event on the contacts of a mechanical relay obtained by measuring with an oscilloscope. A simple measurement circuit for measuring the voltage at the relay contact is shown in Figure 9.

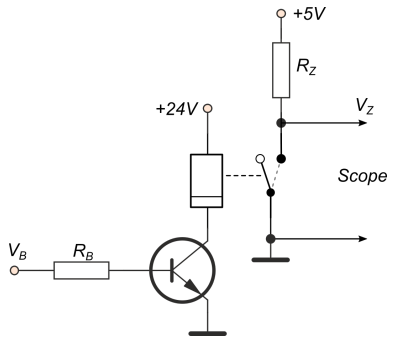


Figure 9: Measurement of electrical properties of mechanical relay contact

Due to inertia and mechanical resonance, the contact generates a delayed series of short pulses that can negatively affect the operation of the device that evaluates them. The waveform measured by the oscilloscope was saved in a data file in the text format (CSV) which will be read as the input data to the simulation.
The object of the problem is to design protection circuits for the safe connection of a relay mechanical contact to the microcontroller port and to prevent false evaluation of impulses caused by contact oscillations. The microcontroller input port is a separately programmable peripheral enabling direct connection of digital circuits as well as analog peripherals. A simplified microcontroller port circuit diagram is shown in Fig. 11 taken from the manufacturer's documentation. The circuit includes port protection diodes, auxiliary resistors for port operation in PULL-UP and PULL-DOWN mode, input-output circuits for digital circuits and a multiplexer for internal analog peripherals connection. The example is focused on the basic properties of the circuit, more complicated cases for the connection of peripherals by long lines with parasitic inductances, capacitances and interference is the subject of further examples.

The circuit connection including only the essential parts for the digital input circuits is shown in

Figure 10: Relay contact switching, V_B - yellow line, V_Z - blue line

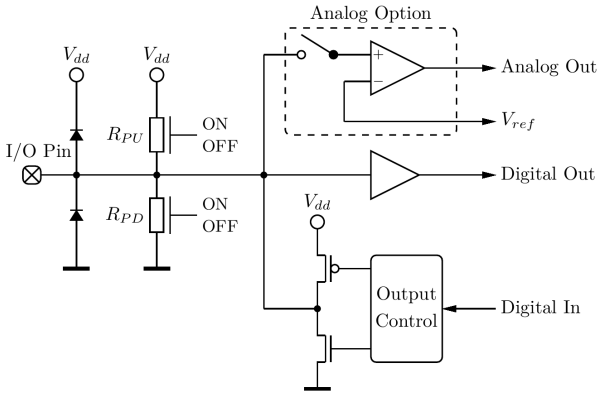


Figure 11: STM32L432 microcontroller port circuit

Fig. 12. The file name containing the input data is an *input* attribute of the component $A1$. The $A2$ component is a model of the comparator which is part of a port with parameters taken from the technical documentation of the microcontroller.

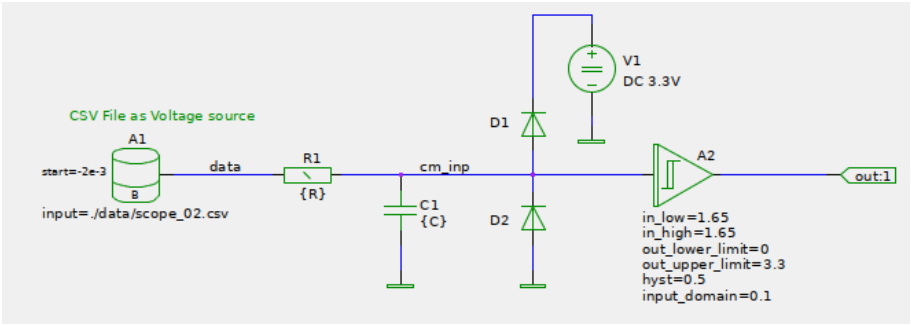


Figure 12: Input circuits

During the simulation, it is possible to change the circuit elements, define component models and monitor the response of the circuit to changes.

```
g = gSim('./examples/port_01.sch')
g.netlist()

g.setPAR('R', '5k')
g.setPAR('C', '10nF')
g.setTRAN(1e-6, 6e-3, 1e-3)

MODEL_1N4007 = 'IS=7.02767e-09 RS=0.0341512 N=1.80803 EG=1.05743 ' + \
               'XTI=5 BV=1000 IBV=5e-08 CJO=1e-11 '+ \
               'VJ=0.7 M=0.5 FC=0.5 TT=1e-07 '+ \
               'KF=0 AF=1'

g.setMODEL('DMODEL', 'D ' + MODEL_1N4007 )
g.sim()
```

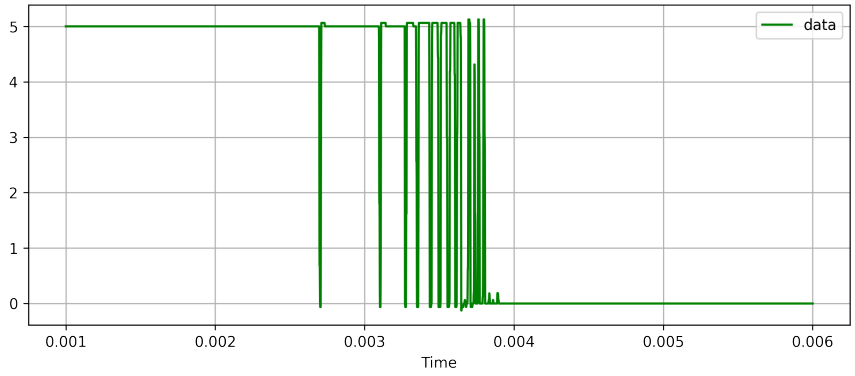


Figure 13: Simulation input data, real signal measured by oscilloscope

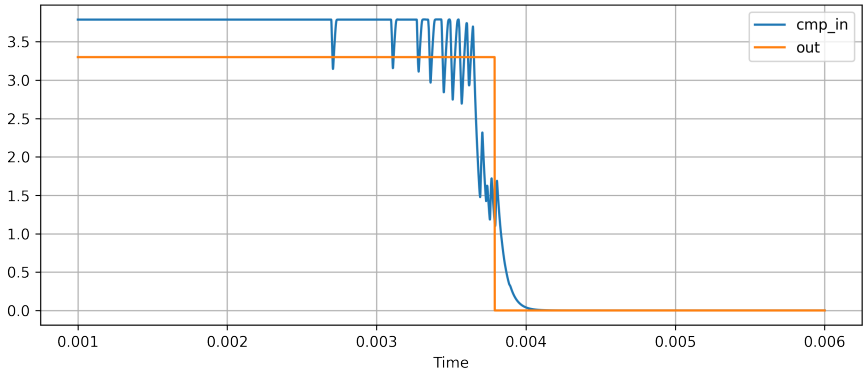


Figure 14: The simulation results, voltages at the comparator A2 input and output

Conclusion

The presented concept demonstrates the possibilities of creating active documents to support the student's independent work within self-study or distance learning. In the process of education this concept allows to concentrate within Jupyter-Notebook all the prerequisites needed to present the topic without disruptive browsing between different programs and environments. The advantage of this concept is that it is based exclusively on open-source software which allows its modification and adjustment for specific use without any restrictions. The current version of the concept is tested in a version for individual use with locally installed software. It is possible to modify it using current technologies (*Jupyter Hub, Docker*) for use within a group of students with global or individual problems assignment and with the extension of possibilities to actively test the topic by solving the problems.

References

[1] Kluyver, T. et al., 2016. Jupyter Notebooks – a publishing format for reproducible computational workflows. In F. Loizides & B. Schmidt, eds. Positioning and Power in Academic Publishing: Players, Agents and Agendas. pp. 87–90.

[2] Barba, A. et al, 2019. Teaching and Learning with Jupyter. https://jupyter4edu.github.io/jupyter-edu-book/

[3] https://executablebooks.org/en/latest/gallery.html

[4] https://jupyterbook.org/intro.html

[5] https://jupyter.org/hub

[6] https://tljh.jupyter.org/en/latest/

[7] https://nbviewer.jupyter.org/

[8] A. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J.D. Joannopoulos, and S.G. Johnson, MEEP: A flexible free-software package for electromagnetic simulations by the FDTD method, Computer Physics Communications, Vol. 181, pp. 687-702, 2010

[9] A. Pursula, M. Lyly, E. Järvinen, T.Zwinger and J.Ruokolainen, Multiphysical simulations with Elmer finite element software in combination with GiD, a paper presented in GiD 2006 seminar in Barcelona, March 2006.

[10] http://bwrcs.eecs.berkeley.edu/Classes/IcBook/SPICE/

[11] http://ngspice.sourceforge.net/

[12] https://github.com/lepton-eda/lepton-eda

[13] Fabo, Peter ; Pavlikova, Sona; (2010). Gsim — Software for simulation in electronics. S4-14. 10.1109/EPEPEMC.2010.5606571, DOI:10.1109/EPEPEMC.2010.5606571

Inverses and generalized inverses of trees: A brief survey

Soňa Pavlíková
Faculty of Chemical and Food Technology, Slovak University of Technology
sona.pavlikova@stuba.sk

Abstract: *The inverse of a labeled graph with a non-singular adjacency matrix is a labeled graph (uniquely determined by the original one up to isomorphism) with spectrum consisting precisely of the reciprocals of the eigenvalues of the original graph. We survey a selection of results on inverses of labeled trees, introduce briefly various approaches to 'inverting' non-invertible matrices, review a formula for a generalized inverse of a labeled tree, and outline main ideas furnishing a new proof of this formula.*

Keywords: inverses of graphs, Moore - Penrose inverse, adjacency matrix, labeled graph

INTRODUCTION

A considerable part of research into spectra of graphs is motivated by application in chemistry. For example, the smallest positive and the largest negative eigenvalue of a graph representing a molecule play an important role in quantum chemistry by determining the minimal binding energy of the molecule. In somewhat more detail, energy of the highest occupied molecular orbital (HOMO) and of the lowest unoccupied molecular orbital (LUMO) of the molecule correspond, respectively, to the smallest positive and the largest negative eigenvalue of the graph; their difference is the well known HOMO-LUMO separation gap.

Lower bounds on the smallest positive eigenvalue of a graph are available but they are not as abundant as the upper bounds on the spectral radius. If, however, one could 'invert' a graph in the sense of inverting entries in its spectrum, then any upper bound on the largest eigenvalue of an 'inverse' graph would automatically be a lower bound for the smallest positive eigenvalue of the original graph. This is one of the main application motivation for the study of 'inverses' of graphs, some of which we now briefly review.

A straightforward way of thinking of an inverse of a (multi-)graph would be to invert its adjacency matrix in the case it is non-singular. It turns out, however [5], that in such a case the inverted matrix has non-negative integral entries if and only if the graph is a union of isolated edges. Another approach could consist in considering an inverse of a graph to be any graph the spectrum of which is obtained by inverting every non-zero eigenvalue (including multiplicities) of the original graph. Since every symmetric matrix is diagonalizable, the above is equivalent to defining a graph H to be an inverse of a graph G if the adjacency matrix A_H of H is similar to the inverse of the adjacency matrix A_G of G. As entries of A_G are non-negative integers, such a definition would imply that $\det(A_G) = \pm 1$ and hence if H is in this sense an inverse of G, entries of A_H would be integral. Such a way of proceeding proved fruitful in a number of ways [4] but suffers from the aesthetical drawback that inverses, if they exist, would not be unique in general. A way out is to restrict similarity to signability and declare a graph H to be an inverse of G if $A_H = DA_G^{-1}D$ for some diagonal ± 1 matrix D. This was first suggested in [4] and later elaborated in [6]; this setting also implies the desirable relation $(G^{-1})^{-1} = G$ if the inverse exists. For more information about the history of investigation of graph inverses we also recommend [6].

The main problem with existence of inverse in any of the above sense is the fact that for 'most' (multi-)graphs with no zero eigenvalue the inverse of their adjacency matrix is not signable (or, even more generally, not similar) to a matrix with *non-negative* entries. Most of the research in [4, 6] therefore focused on sufficient conditions for a graph G with no zero eigenvalue to have A_G^{-1} similar or signable to a *non-negative* matrix. To arrive at a reasonably large set of 'invertible graphs' it makes sense to consider labelled graphs. As edge multiplicity can be expressed by means of an appropriate label, from now on we will consider *simple* graphs, that is, having neither multiple edges nor loops.

Let G be a simple graph with edge set E_G and let $\mathcal{K}$ be a (not necessarily commutative) ring. A *labeling* $\alpha : E_G \to \mathcal{K}$ is an arbitrary function assigning to every edge $e \in E_G$ a non-zero label $\alpha(e) \in \mathcal{K}$. The pair (G, α) is a *labeled graph*; the ring $\mathcal{K}$ does not appear in the notation and will always be understood from the context. As usual, an adjacency matrix $A_{(G,\alpha)}$ is a square matrix with rows and columns indexed by the vertex set of G, in which the uv-th element a_{uv} is equal to zero if u and v are not adjacent in G, and $a_{uv} = \alpha(e) \neq 0$ if u and v are joined by the edge e.

If $A_{(G,\alpha)}$ is non-singular, we define an *inverse* of a labeled graph (G, α) to be a labeled graph (H, β) with labels in the same ring $\mathcal{K}$ and with adjacency matrix $A_{(H,\beta)} = A_{(G,\alpha)}^{-1}$.

In this context, however, one may ask what can be done in the case of labeled graphs with a *singular* adjacency matrix. A most straightforward approach is to invoke one of the (many) generalizations of matrix 'inverses', such as the ones due to Moore-Penrose or Drazin (a special case of the latter being known as the group inverse). A basic question that arises then is whether or not a 'generalized inverse' of a graph can be derived solely from the structure of the graph.

The purpose of this short survey is to review fundamental facts about 'classical' graphs, with focus on inverses of trees with a non-singular adjacency matrix, and to address the questions stated above for 'generalized' inverses of trees, allowing also those with singular adjacency matrix.

1 Inverses of labeled graphs

It is well known that an adjacency matrix of a tree is invertible if and only if the tree contains a perfect matching (which is then necessarily unique). It therefore makes sense to consider a wider class of bipartite graphs with a unique perfect matching and look for properties which would enable for a description of inverses based just on the structure of the graph, without actually inverting its adjacency matrix.

Thus, let (G, α) be a labeled graph, with edge labels in a (not necessarily commutative) ring $\mathcal{K}$, such that G is bipartite and has a unique perfect matching. Since G is bipartite, the adjacency matrix $A_{(G,\alpha)}$ may be assumed to have the block form

$$A_{(G,\alpha)} = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} ; \tag{1}$$

here A is usually called a *bipartition matrix* of (G, α). By [4, Lemma 2.1] we know that a simple bipartite graph G as above has a unique perfect matching if and only if its vertex set V_G admits a bipartition such that vertices in both parts can be linearly ordered in such a way that the above bipartition matrix A is triangular; we will assume this from now on. Then, the matrices A and $A_{(G,\alpha)}$ are invertible if and only if all diagonal entries of A have multiplicative inverses in $\mathcal{K}$.

If a bipartite graph G with a unique perfect matching M has $2n$ vertices and its bipartition matrix A is upper triangular, then we may (and we will) without loss of generality assume that the bipartition of V_G has the form $\{1, 2, \ldots, n\}$ and $\{1', 2', \ldots, n'\}$, with ii' being the edges of M and with no edge in G of the form ij' for $i > j$, where $i, j \in \{1, 2, \ldots, n\}$. We will briefly refer to this situation by saying that both G and A are in an *upper canonical form*. A *lower canonical form* of G and A is defined analogously. The matched and unmatched edges of G will occasionally be called *horizontal* and *descending*, respectively, as they can be drawn as horizontal and descending segments (from left to right) when the non-dashed and dashed vertices are drawn in two columns next to each other in an obvious way. We will use this notation and terminology throughout from this point on. Note that if (H, β) is the inverse of our labeled graph (G, α) with G in an upper canonical form, then H is automatically represented in a lower canonical form.

By an $u \to v$ path P in G we understand a sequence $u_0 u_1 \ldots u_\ell$ of mutually distinct vertices of G with $u_0 = u$, $u_\ell = v$, and $u_{k-1}u_k \in E_G$ for every $k \in \{1, \ldots, \ell\}$. Such a path P will be called *M-alternating*, or simply *alternating*, if ℓ is odd and $u_{k-1}u_k \in M$ if and only if k is odd, $1 \leq k \leq \ell$. We will say that an alternating path P is *even* (*odd*) if it contains and even (odd) number of edges not in M. Thus, if P consists of a single edge $e \in M$, then P is even. Letting $\alpha_k = \alpha(u_{k-1}u_k)$ and recalling that labels of edges of M are assumed to have multiplicative inverses in $\mathcal{K}$, for an alternating $u_0 \to u_\ell$ path P in (G, α) as above we define the *value* $\omega_\alpha(P)$ of P to be

$$\omega_\alpha(P) = \alpha_1^{-1}\alpha_2\alpha_3^{-1}\alpha_4 \ldots \alpha_{\ell-2}^{-1}\alpha_{\ell-1}\alpha_\ell^{-1} \ . \tag{2}$$

That is, to obtain the value $\omega_\alpha(P)$ we multiply through the inverses of labels of matched edges and the original labels of unmatched edges in the order the path is traversed. Finally, for a pair of distinct vertices u, v of G we let $p_M^+(u, v)$ and $p_M^-(u, v)$ be the sum of the *values* $\omega_\alpha(P)$ of all even and odd alternating $u \to v$ paths P, respectively. Note that the values of $p_M^+(u, v)$ and $p_M^-(u, v)$ are automatically zero if both $u, v \in \{1, 2, \ldots, n\}$ or both $u, v \in \{1', 2', \ldots, n'\}$. Observe also that, for $i \leq j$, alternating $i' \to j$ paths in our graph G will always have the form $i'ir'rs's \ldots t'tj'j$ for some (possibly empty, if $i = j$) set of vertices $r, s, \ldots, t$ such that $i < r < s < \ldots < t < j$.

Generalizing the ideas contained in the (unpublished) PhD dissertation of the author [8] we prove that the labeled graphs considered above automatically have inverses, following the original outline given in [8].

Theorem 1 [9] *Let G be a simple bipartite graph of order $2n$ in am upper canonical form with a unique perfect matching M and let $\alpha : E_G \to \mathcal{K}$ be a labeling in a (not necessarily commutative) ring $\mathcal{K}$ such that the label of every edge in M has a multiplicative inverse in $\mathcal{K}$. Then, the labeled graph (G, α) has an inverse (H, β) whose lower canonical form on the vertex set $V_H = V_G$ is given by letting two distinct vertices $i' \in \{1', 2', \ldots, n'\}$ and $j \in \{1, 2, \ldots, n\}$, $i \leq j$, be adjacent in H if and only if $p_M^+(i', j) \neq p_M^-(i', j)$, and by defining $\beta(i'j) = p_M^+(i', j) - p_M^-(i', j)$ for $i'j \in E_H$.*

Let (H, β) be the inverse of (G, α) as in Theorem 1. The graphs G and H have the same vertex set but their edge sets have just the edges of the unique perfect matching M in common because of an upper and a lower canonical form of G and H, respectively. We now show that (at least part of) G can be embedded in H. Let G' be the subgraph of G on the same vertex set $V_{G'} = V_G$ with the edge set $E_{G'} = \{i'j;\ ij' \in E_G;\ \beta(i'j) \neq 0\}$. We make G' into a labeled graph (G', α') by letting $\alpha'(i'j) = \beta(i'j)$ and we will call (G', α') the *derived graph* of (G, α). Note that G' is isomorphic to a subgraph of G via the bijection interchanging i with i', $1 \leq i \leq n$. The derived graph (G', α') is a *labeled subgraph* of the inverse (H, β) of (G, α) in the sense that G' is a subgraph of H and the labelings α' and β coincide on edges of G'. Observe also that all edges $e \in M$ appear in both G' and H, with labels $\alpha'(e) = \beta(e) = \alpha(e)^{-1}$. We sum up these facts as follows.

Theorem 2 [9] *Let G be a simple bipartite graph with a unique perfect matching and with a labeling in a ring $\mathcal{K}$ assigning to every matched edge an invertible label, and let (G, α) be a labeled graph in an upper canonical form. Then, the derived graph (G', α') is a labeled subgraph of the inverse of (G, α).*

Suppose, for example, that the underlying graph G of our labeled graph (G, α) considered above is a tree T. For every unmatched edge $ij' \in E_T$ $(i < j)$ we then have a unique (and, as it happens, odd) alternating $i' \to j$ path $P = i'ij'j$ and so $p_M^+(i', j) - p_M^-(i', j) = -\omega_\alpha(P) = -\alpha(ii')^{-1}\alpha(ij')\alpha(jj')^{-1} \neq 0$. It follows that $i'j$ is an edge of the derived graph (T', α'), so that T' can be identified with T, and $\alpha'(i'j) = -\alpha(ii')^{-1}\alpha(ij')\alpha(jj')^{-1}$. Of course, for every matched edge ii' of $T = T'$ we have $\alpha'(ii') = \alpha(ii')^{-1} = a_{ii'}^{-1}$. Lemma 2 then shows that the inverse of a labeled tree can be considered to be a super-graph of the tree. We state these observations for a later use.

Theorem 3 [9] *Let T be a tree of order $2n$ with a unique perfect matching and let (T, α) be a labeled graph, given in an upper canonical form. The derived graph (T', α') has T' isomorphic to T, with $\alpha'(ii') = \alpha(ii')^{-1}$ for every matched edge and $\alpha'(i'j) = -\alpha(ii')^{-1}\alpha(ij')\alpha(jj')^{-1}$ for every unmatched edge $(i < j)$ of T. Moreover, (T', α') is a labeled subgraph of the inverse of (T, α).*

2 Generalized inverses of trees

As alluded to in the Introduction, it is natural to ask what one can do in the case of labeled graphs with a *singular* adjacency matrix. An equally natural move is to consider 'inverting' the matrix by taking one of the generalizations of matrix inverses, such as the Moore-Penrose inverse, or the Drazin inverse, or a special case of the latter known as the group inverse. In the instance of a (square) symmetric matrix A all these inverses coincide and are commonly called a *pseudo-inverse* of A, which we will denote by A^* throughout. The pseudo-inverse of a *symmetric* matrix is easy to introduce as follows. Since a real symmetric $n \times n$ matrix A is orthogonally diagonalizable, there is an orthogonal matrix P such that $PAP^T = D$, where $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_k, 0, \ldots, 0)$ is the diagonal matrix of eigenvalues of A, with $k = \mathrm{rank}(A)$ non-zero eigenvalues $\lambda_1, \ldots, \lambda_k$. Letting $D^* = \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_k^{-1}, 0, \ldots, 0)$, the *pseudo-inverse* A^* of A is simply given by $A^* = PD^*P^T$, that is, both A and A^* are conjugate to their corresponding diagonal matrices by the *same* orthogonal matrix P. Note that A^* is again symmetric, and A^* coincides with A^{-1} if A is non-singular.

Motivated by this, we define the *pseudo-inverse* of a labeled graph (G, a) with adjacency matrix A to be the labeled graph (G^*, a^*) with adjacency matrix A^*, the pseudo-inverse of A. As before, G and G^* are assumed to have the same vertex set, and $e = uv$ is an edge of G^* if and only if the uv-th entry of A^* is non-zero, and then this entry is also the weight $a^*(e)$ of e. And, again, note that G^* is well defined up to isomorphism preserving edge weights.

Observe that this way of defining pseudo-inverses of labeled graphs is in line with the original motivation of considering graph inverses which comes from chemistry. Namely, there appear to be fewer methods for estimating the smallest positive eigenvalue of a graph in contrast to a larger number of techniques for bounding the largest positive eigenvalue. For graphs representing structure of molecules, however, the smallest positive eigenvalue is a

meaningful parameter in quantum chemistry. If such a graph has an inverse, one may hope to increase the number of techniques for estimating its smallest positive eigenvalue by passing to bounds on the largest positive eigenvalue of the inverse graph. This feature remains present also for our pseudo-inverses.

A formula for entries of the adjacency matrix of the pseudo-inverse of a tree with arbitrary non-zero edge weights can be derived from a result of [3] stated in terms of bipartite graphs associated with arbitrary matrices (with the vertex set being the union of row and column indices of a matrix) in the special case when the graphs are acyclic. The proof of the result of [3] is based on a determinant formula for entries of the Moore-Penrose inverse that first appeared in a classical paper [7]; for more recent references see [1] or [2, Appendix A].

In [11] we gave a different proof of a formula for calculating the pseudo-inverse of an arbitrary labeled *tree*. The proof in [11] does not refer to the formulae for entries of the Moore-Penrose inverse and is based solely on considering eigenvectors (without actual evaluation of any of them) regarded as functions on the vertex set. In fact, some of our considerations are valid for any real-valued function on the vertex set of a tree (not necessarily those representing eigenvectors).

To state the result we need to introduce a few concepts. Let (T, a) be a labeled tree; for brevity we will often omit the symbol for the weight function in our exposition. For an unordered pair of vertices u, v of distinct vertices of T we let $\mathcal{M}(u, v)$ denote the set of all maximum matchings M of T with the property that edges of the (unique) $u-v$ path in T belong alternately to M and not to M, with the condition that both the first and the last edge of the path (that is, those incident to u and v) belong to M. A necessary condition for the set $\mathcal{M}(u, v)$ to be non-empty is that the distance between u and v be odd, but note that this condition does not need to be sufficient; though, if uv is an edge of some maximum matching, then the set $\mathcal{M}(u, v)$ is automatically non-empty. A pair of vertices u, v for which $\mathcal{M}(u, v) \neq \emptyset$ will be called *maximally matchable*.

Further, for any maximally matchable pair of vertices u, v and a maximum matching $M \in \mathcal{M}(u, v)$ let $\alpha_{\overline{u,v}}(M)$ denote the product of all the weights $a(e)$, ranging over all edges e of M that are not contained in the unique $u-v$ path P in T. (The line over the pair of vertices u, v in the subscript indicates that edges of $M \cap P$ are not considered in the product; a product over an empty set is considered to be equal to 1.) Also, for the same pair of vertices u, v let $\alpha(u, v)$ be the product of all the values of $a(e)$ taken over all the

edges e in the path P (necessarily of odd length), and multiplied by $+1$ or -1 depending on whether the distance between u and v is congruent to $+1$ or -1 mod 4; if u, v are not maximally matchable we set $\alpha(u, v) = 0$. With this in hand we may associate with any maximally matchable pair of vertices u, v of T the value

$$\mu_T(u, v) = \alpha(u, v) \cdot \sum_{M \in \mathcal{M}(u,v)} (\alpha_{\overline{u,v}}(M))^2 \ ;$$

it follows that $\mu_T(u, v) = 0$ if u, v is not a maximally matchable pair (which includes the case $u = v$). Finally, letting $\mathcal{M}$ be the set of all maximum matchings in T, for every $M \in \mathcal{M}$ let $\alpha(M)$ be the product of the weights $a(e)$ taken over all edges e of M, and let

$$m(T) = \sum_{M \in \mathcal{M}} (\alpha(M))^2 \ .$$

In this terminology and notation we have:

Theorem 4 [11] *Let (T, a) be a labeled tree with vertex set V. Then, its pseudo-inverse (T^*, a^*) has two distinct vertices $u, v \in V$ joined by an edge e if and only if u, v is a maximally matchable pair in T, with weight of e given by*

$$a^*(e) = a^*(uv) = \frac{\mu_T(u, v)}{m(T)} \ .$$

It may be of independent interest to sketch the ideas upon which the proof of this theorem in [11] is based. A fundamental observation from elementary linear algebra that we make use of is the following.

Theorem 5 *Two symmetric square matrices A and B of the same dimension are generalized inverses of each other if and only if they have the same null-space and, for every non-zero eigenvalue λ of A the quantity λ^{-1} is an eigenvalue of B and the corresponding eigenspaces of A and B are identical. More explicitly, if $A = (a_{ij})$ and $B = (b_{ij})$, then B is the generalized inverse of A if and only if A and B have the same null-space, and every eigenvector $f : [n] \to \mathbb{R}$ of A corresponding to a non-zero eigenvalue of A satisfies $\sum_{j \in [n]} b_{ij} \sum_{k \in [n]} a_{jk} f(k) = f(i)$* for every $i \in [n]$.

A second ingredient of fundamental importance and of independent interest, used in [11], is the following, which uses the definition of the parameters $\mu_T(u,v)$ and $m(T)$ introduced above, is:

Theorem 6 *Let A be an adjacency matrix of a labeled tree (T,a) with vertex set V and let B be a matrix (indexed the same way as A by elements of V) with uv-th entry equal to $\mu_T(u,v)/m(T)$ for every $u,v \in V$. Then, A and B are Gaussian equivalent.*

The actual proof of Theorem 4 in [11] is then a combination of the preceding two results with further calculations related to alternating paths in trees.

We conclude with and example in which we determine the generalized inverse of the path P_5 on 5 vertices. Here, $m(P_5) = 3$, and in what follows we let A, D and P^T be, respectively. the adjacency matrix of P_5, the diagonal matrix with eigenvalues of P_5, and the orthonormal matrix with columns formed by the corresponding orthogonal eigenvectors P_5.

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{3} & 0 \\ 0 & 0 & 0 & 0 & -\sqrt{3} \end{pmatrix}$$

$$P^T = \begin{pmatrix} \sqrt{3}/3 & 0 & -\sqrt{3}/3 & 0 & \sqrt{3}/3 \\ -1/2 & -1/2 & 0 & 1/2 & 1/2 \\ -1/2 & 1/2 & 0 & -1/2 & 1/2 \\ \sqrt{3}/6 & 1/2 & \sqrt{3}/3 & 1/2 & \sqrt{3}/6 \\ \sqrt{3}/6 & -1/2 & \sqrt{3}/3 & -1/2 & \sqrt{3}/6 \end{pmatrix}$$

Let D^* be formed from D by inverting the non-zero entries. The adjacency matrix A^* of the pseudoinverse P_5^* of the path P_5 then is

$$A^* = PD^*P^T = \begin{pmatrix} 0 & 2/3 & 0 & -1/3 & 0 \\ 2/3 & 0 & 1/3 & 0 & -1/3 \\ 0 & 1/3 & 0 & 1/3 & 0 \\ -1/3 & 0 & 1/3 & 0 & 2/3 \\ 0 & -1/3 & 0 & 2/3 & 0 \end{pmatrix}$$

CONCLUSION

In this paper we gave a brief survey of results on (generalized) inverses of labeled trees, including an outline of a new proof of a formula for these generalized inverses.

A possible new avenue of research in this area would be to investigate generalized inverses of connected graphs which are not trees. A first step in this direction has been done in [10] by considering generalized inverses of cycles.

Acknowledgement The authors acknowledge support of this research by the APVV Research Grants 17-0428 and 19-0308, as well as from the VEGA Research Grants 1/0238/19 and 1/0206/20.

References

[1] A. Ben-Israel, The Moore of the Moore-Penrose inverse, Electron. J. Linear Algebra 9 (2002), 150–157.

[2] A. Ben-Israel, T. N. E. Greville, 'Generalized inverses', Theory and Applications, 2nd ed., Springer, 2003.

[3] T. Britz, D. D. Olesky and P. van den Driessche, The Moore-Penrose inverse of matrices with an acyclic bipartite graph, Linear Algebra Appl. 390 (2004), 47–60.

[4] Godsil, C. D.: Inverses of Trees, Combinatorica 5 (1985), 33–39.

[5] F. Harary, The determinant of the adjacency matrix of a graph, SIAM Review 4 (1962), 202–210.

[6] C. McLeman, and A. McNicholas, Graph invertibility, Graphs Combin. 30 (2014), 977–1002.

[7] E. Moore, On the reciprocal of the general algebraic matrix, Bull. Amer. Math. Soc. 26 (1920), 394–395.

[8] Pavlíková, S.: *Graphs with unique* 1*-factor and matrices*, PhD Dissertation, Comenius University, Bratislava, 1994.

[9] Pavlíková, S.: *A note on inverses of labeled graphs*, Australasian J. Combinatorics 67 (2017), 222–234.

[10] Pavlíková, S., Krivoňaková, N: *Generalized inverses of cycles*, V Mathematics, Information Technologies and Applied Sciences 2018, University of Defence, Brno, (2018).

[11] S. Pavlíková, J. Širáň: Inverting non-invertible trees, *Australasian Journal of Combinatorics*, vol. 75 (2), 2019, 246 -255

Název: Mathematics, Information Technologies and Applied Sciences 2021
Editoři: Jaromír Baštinec
Miroslav Hrubý
Vydavatel: Univerzita obrany, Brno
Rok vydání: 2021

Publikace neprošla jazykovou úpravou.

ISBN 978-80-7582-441-7 (e-kniha)