

INTERNATIONAL SCIENTIFIC JOURNAL

MATHEMATICAL MODELING

YEAR II, ISSUE 3/2018
ISSN (PRINT) 2535-0986
ISSN (WEB) 2603-2929

Published by
SCIENTIFIC-TECHNICAL UNION of MECHANICAL ENGINEERING - INDUSTRY 4.0
Sofia, BULGARIA

INTERNATIONAL EDITORIAL BOARD

EDITOR I CHIEF		
Prof. ANDREY FIRSOV Peter the Great St.Petersburg Polytechnic University		RU

Members:		
Abilmazhin Adamov, Prof.	L.N.Gumilyov Eurasian National University	KZ
Alexander Guts, Prof.	Omsk State University	RU
Alexei Zhabko, Prof.	Saint Petersburg State University	RU
Andrey Markov, Prof.	Baltic State Technical University	RU
Andrii Matviichuk, Prof.	Kyiv National Economics University	UA
Andrzej Nowakowski, Prof.	University of Lodz	PL
Anton Makarov, Dr.	Saint Petersburg State University	RU
Armands Gricans, Assoc. Prof.	Daugavpils University	LV
Artūras Dubickas, Prof.	Vilnius University	LT
Avinir Makarov, Prof.	Saint Petersburg State University of Industrial Technologies and Design	RU
Christo Boyadjiev, Prof.	Institute of Chemical Engineering, BAS	BG
Daniela Marinova, Dsoc. Prof.	Technical University of Sofia	BG
Dimitrios Poulakis, Prof.	Aristotle University of Thessaloniki	GR
Evgeniy Smirnov, Assoc. Prof.	Volgograd State Technical University	RU
Galia Angelova, Prof. DSc,	Bulgarian Academy of Science	BG
Giovanni Borgioli, Assoc. Prof.	University of Florence	IT
Haskiz Coskun, Prof.	Karadeniz Technical University of Trabzon	TR
Idilia Bachkova, Prof.	University of Chemical Technology and Metallurgy	BG
Igor Anufriev, Assoc. Prof.	Peter the Great St.Petersburg Polytechnic University	RU
Irena Stojkovska, Prof.	Ss. Cyril and Methodius University in Skopje	MK
Ivana Štajner-Papuga, Prof.	University of Novi Sad	RS
Kanagat Aldazharov, Assoc. Prof.	Kazakh Economics University	KZ
Karl Kunisch, Prof.	University of Graz	AT
Mahomed Agamirza ogly Dunyamalyev, Prof.	Azerbaijan Technical University	AZ
Marius Giuclea, Prof.	The Bucharest University of Economics Studies	RO
Mihail Okrepilov, Prof.	D.I. Mendeleyev Institute for Metrology (VNIIM)	RU
Milena Racheva, Assoc. Prof.	Technical University of Gabrovo	BG
Mohamed Kara, Dr.	Ferhat Abbas Sétif 1 University	DZ
Mohamed Taher El-mayah, Prof	MTI University	EG
Neli Dimitrova, Prof.	Institute of Mathematics and Informatics, BAS	BG
Nina Bijedic, Prof.	Dzemal Bijedic University of Mostar	BA
Oleg Obradović, Prof.	University of Montenegro	ME
Olga Pritomanova, Assoc. Prof.	Oles Honchar Dnipropetrovsk National University	UA
Özkan Öcalan, Prof.	Akdeniz University of Antalya	TR
Pașc Găvruță, Prof.	Politehnic University of Timisoara	RO
Pavel Satrapa, Assoc. Prof.	Technical University of Liberec	CZ
Pavel Tvrdík, Prof.	Czech Technical University in Prague	CZ
Pavlina Yordanova, Assoc. Prof.	Shumen University	BG
Petr Trusov, Prof.	Perm State Technical University	RU
Rannveig Björnsdóttir, Prof.	University of Akureyri	IS
Roumen Anguelov, Prof.	University of Pretoria	ZA
Sándor Szabó, Dr. Prof.	University of Pécs	HU
Sashko Martinovski, Assoc. Prof.	St. Kliment Ohridski University of Bitola	MK
Sergey Bosnyakov, Prof.	Moscow Institute of Physics and Technology	RU
Sergey Kshevetskii, Prof.	Immanuel Kant Baltic Federal University	RU
Snejana Hristova, Prof.	University of Plovdiv	BG
Svetlana Lebed, Assoc. Prof.	Brest State Technical University	BY
Tomasz Szarek, Prof.	University of Gdansk	PL
Valeriy Serov, Prof.	University of Oulu	FI
Vasily Maximov, Prof.	Saint Petersburg State University of Industrial Technologies and Design	RU
Ventsi Rumchev, Prof.	Curtin University, Perth	AU
Veronika Stoffová, Prof.	University of Trnava	SK
Veselka Pavlova, Prof.	University of National and World Economy	BG
Viorica Sudacevschi, Assoc. Prof.	Technical University of Moldova	MD
Vladimir Janković, Prof.	University of Belgrade	RS
Vladislav Holodnov, Prof.	Saint Petersburg State Institute of Technology	RU
Vyacheslav Demidov, Prof.	Saint Petersburg State University of Industrial Technologies and Design	RU
Yordan Yordanov, Assoc. Prof.	University of Sofia	BG
Yuriy Kuznetsov, Prof.	Nizhny Novgorod State University	RU
Zdenka Kolar - Begović, Prof.	University of Osijek	HR

INTERNATIONAL SCIENTIFIC JOURNAL

MATHEMATICAL MODELING

PUBLISHER:

SCIENTIFIC TECHNICAL UNION OF MECHANICAL ENGINEERING
“INDUSTRY 4.0”

108, Rakovski Str., 1000 Sofia, Bulgaria

tel. (+359 2) 987 72 90,

tel./fax (+359 2) 986 22 40,

office@stumejournals.com

www.stumejournals.com

ISSN (PRINT) 2535-0986

ISSN (WEB) 2603-2929

YEAR II, ISSUE 3 / 2018

EDITOR IN CHIEF:

Prof. D.Sc. ANDREY FIRSOV

Peter the Great St.Petersburg Polytechnic University, Russia

CONTENTS

THEORETICAL FOUNDATIONS AND SPECIFICITY OF MATHEMATICAL MODELLING

SUBSTANTIATION OF J.-G. SUN'S HYPOTHESIS, WHICH LIES IN THE BASIS OF THE THEORY OF ANALYTICAL DEPENDENCE OF EIGENVALUES OF MATRIX FROM "DISTURBING" PARAMETERS UNDER MULTI-PARAMETRIC PERTURBATION OF THE MATRIX ELEMENTS

Prof. Dr.Tech.Sci. Andrei N. Firsov, Master's student Le Van Khanh 88

THE COMPUTATION OF ROTOR'S MOTION IN CYLINDRICAL CHAMBER FILLED WITH VISCOUS GAS

Prof., Dr. Dementev O. 91

IMPLICIT EULER TIME DISCRETIZATION AND FDM WITH NEWTON METHOD IN NONLINEAR HEAT TRANSFER MODELING

Ph.D. Filipov S., Prof. D.Sc. Faragó I. 94

MATHEMATICAL MODELLING OF TECHNOLOGICAL PROCESSES AND SYSTEMS

MATHEMATICAL MODELS OF CALCULATIONS OF PARAMETERS OF CRYSTALLIZATION OF BINARY ALLOYS BY MEANS OF COMPUTER THERMAL ANALYSIS

Ass. Prof., Dr. Eng. Donii O., Ass. Prof., Dr. Eng. Khristenko V., Omelko L., Ass. Kotliar S. 99

MODELING OF MAGNETIC HYSTERESIS UNDER WEAK MAGNETIC FIELDS AND TRIAXIAL STRESS STATE

Mushnikov A.N., Kryucheva K. D. 102

MATHEMATICAL MODELLING OF PIEZOELECTRIC DISK TRANSFORMER WITH RING ELECTRODES IN PRIMARY ELECTRICAL CIRCUIT

Dr. Bazilo C., PhD. 105

MATHEMATICAL MODELLING OF SOCIO-ECONOMIC PROCESSES AND SYSTEMS

USE OF ICT RESOURCES IN THE HUMANITARIAN SUBJECTS IN BULGARIAN SCHOOLS

M.Sc. Boneva Y., Assist. Prof. PhD. Paunova-Hubenova E., Assist. Prof. Terzieva V., PhD. Dimitrov S. 108

COMPARISON OF APPROACHES TO ESTIMATION OF TRANSITION MATRIX FOR THE TERRORIST THREAT MARKOV MODEL

Ing. Martin Tejkal, Doc. RNDr. Jaroslav Michálek, CSc. 112

MATHEMATICAL MODELLING OF MEDICAL-BIOLOGICAL PROCESSES AND SYSTEMS

CLASSIFICATION OF PROTEIN STRUCTURES BY USING FUZZY KNN CLASSIFIER AND PROTEIN VOXEL-BASED DESCRIPTOR

Prof. Dr. Mirceva G., Prof. Dr. Naumoski A., Prof. Dr. Kulakov A. 116

DETERMINATION OF SURFACTANT EFFECTIVE DIFFUSION COEFFICIENT USING INVERSE PROBLEM FOR WARD-TORDAI EQUATION

M.Sc. Eng. Stasiak M., M.Sc. Eng. Andrzejewski A., Prof. Prochaska K. 119

SUBSTANTIATION OF J.-G. SUN'S HYPOTHESIS, WHICH LIES IN THE BASIS OF THE THEORY OF ANALYTICAL DEPENDENCE OF EIGENVALUES OF MATRIX FROM "DISTURBING" PARAMETERS UNDER MULTI-PARAMETRIC PERTURBATION OF THE MATRIX ELEMENTS

ОБОСНОВАНИЕ ГИПОТЕЗЫ J.-G. SUN'А, ЛЕЖАЩЕЙ В ОСНОВЕ ТЕОРИИ АНАЛИТИЧЕСКОЙ ЗАВИСИМОСТИ СОБСТВЕННЫХ ЗНАЧЕНИЙ МАТРИЦЫ ОТ "ВОЗМУЩАЮЩИХ" ПАРАМЕТРОВ ПРИ МНОГОПАРАМЕТРИЧЕСКОМ ВОЗМУЩЕНИИ МАТРИЧНЫХ ЭЛЕМЕНТОВ

Prof. Dr.Tech.Sci. Andrei N. Firsov¹, Master's student Le Van Khanh²
Institute of Computer Science and Technology – Peter the Great Saint-Petersburg Polytechnic University, Russia
E-mail: ¹anfirs@yandex.ru, ²levankhanhth@gmail.com

Abstract: The paper gives a rigorous justification of Ji-Guang Sun's hypothesis about the properties of the eigenvalues of the matrix of a linear dynamical system under multiparametric perturbation of its elements.

KEYWORDS: DYNAMICAL SYSTEMS WITH SMALL PERTURBATIONS, INVERSE STABILITY PROBLEM.

1. Introduction

The concept of stability of dynamic systems characterizes the property of a system to operate stably in modes where there are uncertainties in the values of certain system elements. Due to the complexity, and sometimes the practical impossibility to indicate the necessary and sufficient allowable ranges of variation of the corresponding parameters, at least sufficient estimates may be of great interest. On the other hand, practice shows that the desire for universality of theoretical results, as a rule, leads to great difficulties in the application of such results for solving specific problems. We believe that the consideration of such considerations should underlie the construction of theoretical structures aimed at solving specific applied problems. In this paper, we propose a solution to the problem of conditions sufficient to preserve the stability property of a linear dynamical system under small perturbations of its matrix. In this case, an estimate of the smallness of the perturbation parameters is given.

2. Formulation of the problem

Let's consider a dynamic system:

$$\frac{dX(t)}{dt} = AX(t) \quad (2.1)$$

$$A = (a_{ij})_{i,j=1}^n, X(t) = (x_1(t), x_2(t), \dots, x_n(t))^T.$$

Here A denotes an "unperturbed" matrix. Suppose the matrix A has simple different eigenvalues $\lambda_j, j = 1, 2, \dots, n$ and $\operatorname{Re}(\lambda_j) < 0, j = 1, 2, \dots, n$, that is, the system (2.1) is supposed to be sustainable.

Let further

$$\vec{\varepsilon} = (\varepsilon_{11}, \dots, \varepsilon_{1n}, \varepsilon_{21}, \dots, \varepsilon_{2n}, \dots, \varepsilon_{nn})^T \in \square^{n^2},$$

where ε_{ij} – unknown "perturbations" of the elements of the original matrix. Let's consider perturbation parameters ε_{ij} will be small enough in the sense that the quantity ε_{ij}^2 can be neglected in comparison with ε_{ij} : $\varepsilon_{ij}^2 \square |\varepsilon_{ij}|$. The ultimate goal is a "sufficient" estimate of the magnitude of the elements ε_{ij} of a "perturbed" system

$$\frac{dX(t)}{dt} = A(\vec{\varepsilon})X(t), \quad (2.2)$$

under which the stability property of its solutions is preserved.

Further we will consider that

$$A(\varepsilon_{11}, \varepsilon_{12}, \dots, \varepsilon_{1n}, \varepsilon_{21}, \dots, \varepsilon_{2n}, \dots, \varepsilon_{nn}) = A + \sum_{i,j=1}^n \varepsilon_{ij} A_{ij}, \quad (2.3)$$

where A_{ij} – are known matrices. Ji-Guang Sun has proved [1], that in the case of the validity of "Sun's hypothesis" (see (3.1), (3.2) below), eigenvalues $\lambda_j(\vec{\varepsilon}), j = 1, 2, \dots, n$ of a perturbed matrix $A(\vec{\varepsilon})$ are fairly smooth functions of parameters $\varepsilon_{ij}, i, j = 1, 2, \dots, n$ in the neighborhood of zero. This allows, if we assume the validity of the Sun's hypothesis, to answer the question raised above about the stability conditions of a perturbed system (2.2) (see [2]). In this paper we will show that the "hypothesis" (3.1), (3.2) is in fact a theorem.

Under the norms of matrices and vectors, we will further understand the corresponding Euclidean norms.

3. Proof of Sun's hypothesis

The main results on the dependence of eigenvalues on perturbing parameters were obtained by T. Kato [3] (for one parameter) and by J.-G. Sun [1] (for several parameters). However, when formulating the results, J.-G. Sun expressed a hypothesis about the nature of the dependence of the eigenvalues on the perturbation parameters, which he did not prove. In this paper, we justify this hypothesis.

In [1] J.-G. Sun stated the following hypothesis, on which the proof of the main theorem in the paper [1] was based:

Sun's hypothesis. Let λ_s – non-multiple eigenvalue, generally speaking, asymmetric matrix $A \in \square^{n \times n}$, $\vec{x}_s$ and $\vec{y}_s$ – corresponding right and left eigenvectors, at the same time we accept, that $\|\vec{x}_s\| = 1$ (where $\|\vec{x}_s\| = \sqrt{\sum_{i=1}^n (x_{si})^2}$, $x_{si}, i = 1, \dots, n$ are coordinates of the vector $\vec{x}_s$) and $\vec{y}_s^T \cdot \vec{x}_s = 1$. Then for any λ_s there exist such $\tilde{X}, \tilde{Y} \in \square^{n \times (n-1)}$, that for matrices $X = (\vec{x}_s, \tilde{X})$ and $Y = (\tilde{y}_s, \tilde{Y})$, the following relations hold:

$$Y^T X = I_n, \quad (3.1)$$

$$Y^T A X = \begin{pmatrix} \lambda_s & 0_{1 \times (n-1)} \\ 0_{(n-1) \times 1} & \tilde{A} \end{pmatrix}, \lambda_s \notin \lambda(\tilde{A}). \quad (3.2)$$

where $\lambda_s, s = \overline{1, n}$ – are non-multiple matrix eigenvalues A .

Here we give a detailed *proof* of this result. Let us consider the matrix

$$Z' = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \end{pmatrix}_{n \times n} = \begin{pmatrix} x_{s1} & 1 & 0 & \dots & 0 \\ x_{s2} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{s(n-1)} & 0 & 0 & \dots & 1 \\ x_{sn} & 0 & 0 & \dots & 0 \end{pmatrix}_{n \times n}.$$

Applying the Gram-Schmidt orthogonalization (see, for example [4, 5]) to the matrix Z' , we obtain the matrix $Z = (\tilde{x}_s, \tilde{Y})$ - orthonormal in the columns, i.e.

$$\tilde{Y}^T \tilde{x}_s = 0_{(n-1) \times 1}. \quad (3.3)$$

It is also easy to check that $\det Y = (y_s, \tilde{Y}) \neq 0$.

Let $H = \{h_{i,j}\}_{i=1,j=1}^n \equiv Y^{-1} = (y_s, \tilde{Y})^{-1}$, and consider the matrix $X = (\tilde{x}_s, \tilde{X})$, where

$$\tilde{X}_{(n-1)} = \left[\{h_{i,j}\}_{i=2,j=1}^n \right]^T = \left[(h_2)^T \quad (h_3)^T \quad \dots \quad (h_n)^T \right],$$

where $h_i, i = 2, \dots, n$ - are rows of the matrix H , i.e.

$$\tilde{y}_s^T \tilde{X} = 0_{1 \times (n-1)} \text{ and } \tilde{Y}^T \tilde{X} = I_{(n-1)}. \quad (3.4)$$

From (3.3) and (3.4) we have:

$$\begin{aligned} Y^T X &= (\tilde{y}_s, \tilde{Y})^T (\tilde{x}_s, \tilde{X}) = \begin{pmatrix} \tilde{y}_s^T \tilde{x}_s & \tilde{y}_s^T \tilde{X} \\ \tilde{Y}^T \tilde{x}_s & \tilde{Y}^T \tilde{X} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0_{1 \times (n-1)} \\ 0_{(n-1) \times 1} & I_{n-1} \end{pmatrix} = I_n, \end{aligned}$$

therefore, we have a relationship (3.1).

Consider further:

$$\begin{aligned} Y^T A &= (\tilde{y}_s, Y_2, Y_3, \dots, Y_n)^T (\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n) = \\ &= \begin{pmatrix} \tilde{y}_s^T \tilde{a}_1 & \tilde{y}_s^T \tilde{a}_2 & \dots & \tilde{y}_s^T \tilde{a}_n \\ A^* \end{pmatrix} = \begin{pmatrix} \tilde{y}_s^T A \\ A^* \end{pmatrix} = \begin{pmatrix} \lambda_s \tilde{y}_s^T \\ A^* \end{pmatrix}, \end{aligned} \quad (3.5)$$

where $\tilde{a}_j, j = 1, 2, \dots, n$ are matrix columns of the matrix A , and matrix $A^* \in \mathbb{R}^{(n-1) \times n}$. $Y_j, j = 2, 3, \dots, n$ - are matrix columns of Y .

Next, multiply the expression (3.5) by the matrix X on the right:

$$\begin{aligned} Y^T A X &= \begin{pmatrix} \lambda_s \tilde{y}_s^T \\ A^* \end{pmatrix} (\tilde{x}_s, X_2, X_3, \dots, X_n) = \\ &= \begin{pmatrix} \lambda_s \tilde{y}_s^T \tilde{x}_s & \lambda_s \tilde{y}_s^T X_2 & \dots & \lambda_s \tilde{y}_s^T X_n \\ A^{**} \end{pmatrix} = \\ &= \begin{pmatrix} \lambda_s & 0_{1 \times (n-1)} \\ A^{**} \end{pmatrix}, \end{aligned} \quad (3.6)$$

where $X_j, j = 2, 3, \dots, n$ - are matrix columns of X , and matrix $A^{**} \in \mathbb{R}^{(n-1) \times n}$.

Next we have for AX :

$$AX = \begin{pmatrix} \tilde{a}^1 \\ \tilde{a}^2 \\ \vdots \\ \tilde{a}^n \end{pmatrix} \cdot (\tilde{x}_s, X_2, X_3, \dots, X_n) = (\lambda_s \tilde{x}_s, A'), \quad (3.7)$$

where $\tilde{a}^i$ - are rows of matrix A ; $X_j, j = 2, 3, \dots, n$ - are columns of matrix X , and matrix $A' \in \mathbb{R}^{n \times (n-1)}$.

Next, multiply the expression (3.7) by Y^T on the left:

$$\begin{aligned} Y^T A X &= (\tilde{y}_s, Y_2, Y_3, \dots, Y_n)^T \cdot (\lambda_s \tilde{x}_s, A') \\ &= \begin{pmatrix} \lambda_s & A'' \end{pmatrix}_{(n-1) \times 1}, \end{aligned} \quad (3.8)$$

where $Y_j, j = 2, 3, \dots, n$ - are the columns of matrix Y , and matrix $A'' \in \mathbb{R}^{n \times (n-1)}$.

From the property of matrix associativity, and taking into account expressions (3.6), (3.8), we are convinced of the validity of equality (3.2).

We now give an example of the application of the results.

Let the matrix A have three different negative eigenvalues

$$\lambda_1, \lambda_2, \lambda_3.$$

Let $\tilde{x}_1 = (x_{11} \ x_{12} \ x_{13})^T$ and $\tilde{y}_1 = (y_{11} \ y_{12} \ y_{13})^T$ - right and left eigenvectors corresponding to λ_1 . At the same time, the eigenvectors satisfy the following conditions:

$$\|\tilde{x}_1\| = 1 \text{ и } \tilde{y}_1^T \tilde{x}_1 = 1.$$

The case $x_{11} \neq \pm 1$.

Let construct matrices Y, X in this case (see above). Let

$$Z' = \begin{pmatrix} 1 & 0 \\ \tilde{x}_1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \equiv \begin{pmatrix} x_{11} & 1 & 0 \\ x_{12} & 0 & 1 \\ x_{13} & 0 & 0 \end{pmatrix}, \quad \tilde{x}_1 = \begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \end{pmatrix}.$$

Applying Gram-Schmidt orthogonalization to the matrix Z' we obtain the matrix orthonormal in the columns

$$Z = (\tilde{x}_1, \tilde{Y}) = \begin{pmatrix} x_{11} & \frac{1-x_{11}^2}{\sqrt{1-x_{11}^2}} & 0 \\ x_{12} & \frac{-x_{11}x_{12}}{\sqrt{1-x_{11}^2}} & \frac{x_{13}}{\sqrt{1-x_{11}^2}} \\ x_{13} & \frac{-x_{11}x_{13}}{\sqrt{1-x_{11}^2}} & \frac{-x_{12}}{\sqrt{1-x_{11}^2}} \end{pmatrix},$$

with the condition

$$x_{11} \neq \pm 1. \quad (3.9)$$

Similarly, in accordance with the above, we obtain the matrix $Y = (\tilde{y}_1, \tilde{Y})$:

$$Y = \begin{pmatrix} y_{11} & \frac{1-y_{11}^2}{\sqrt{1-y_{11}^2}} & 0 \\ y_{12} & \frac{-y_{11}y_{12}}{\sqrt{1-y_{11}^2}} & \frac{y_{13}}{\sqrt{1-y_{11}^2}} \\ y_{13} & \frac{-y_{11}y_{13}}{\sqrt{1-y_{11}^2}} & \frac{-y_{12}}{\sqrt{1-y_{11}^2}} \end{pmatrix}. \quad (3.10)$$

Its determinant is equal to

$$\det(Y) = \frac{(1-y_{11}^2)(x_{11}y_{11} + x_{12}y_{12} + x_{13}y_{13})}{(1-y_{11}^2)} = 1,$$

whence it follows that there exists an inverse matrix for Y :

$$H = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ \frac{x_{12}y_{12} + x_{13}y_{13}}{\sqrt{1-y_{11}^2}} & \frac{-y_{11}x_{12}}{\sqrt{1-y_{11}^2}} & \frac{-x_{13}y_{11}}{\sqrt{1-y_{11}^2}} \\ \frac{x_{11}(x_{12}y_{13} - x_{13}y_{12})}{\sqrt{1-y_{11}^2}} & \frac{-x_{11}^2y_{13} - x_{11}y_{11}x_{13} - y_{13}}{\sqrt{1-y_{11}^2}} & \frac{x_{11}^2y_{12} - x_{11}y_{11}x_{12} - y_{12}}{\sqrt{1-y_{11}^2}} \end{pmatrix}$$

Next, we build the matrix $X = (\tilde{x}_1, (h_2)^T, (h_3)^T)^T$, where

h_2, h_3 – rows of matrix H , i.e.

$$X = (\tilde{x}_1, \tilde{X}) = \begin{pmatrix} x_{11} & \frac{x_{12}y_{12} + x_{13}y_{13}}{\sqrt{1-x_{11}^2}} & \frac{x_{11}(x_{12}y_{13} - x_{13}y_{12})}{\sqrt{1-x_{11}^2}} \\ x_{12} & \frac{-y_{11}x_{12}}{\sqrt{1-x_{11}^2}} & -\frac{x_{11}^2y_{13} - x_{11}y_{11}x_{13} - y_{13}}{\sqrt{1-x_{11}^2}} \\ x_{13} & \frac{-x_{13}y_{11}}{\sqrt{1-x_{11}^2}} & \frac{x_{11}y_{12} - x_{11}y_{11}x_{12} - y_{12}}{\sqrt{1-x_{11}^2}} \end{pmatrix}, \quad (3.11)$$

$$\det(X) = 1$$

It is not difficult to see that $X = (Y^{-1})^T$, so $Y^T X = I$, i.e. (3.1) is right.

From (3.10) and (3.11) we finally get:

$$Y^T AX = \begin{pmatrix} \lambda_1 \tilde{y}_1^T \tilde{x}_1 & 0 & 0 \\ 0 & \tilde{a}_{22} & \tilde{a}_{23} \\ 0 & \tilde{a}_{32} & \tilde{a}_{33} \end{pmatrix} \quad (3.12)$$

and $\tilde{A} = \begin{pmatrix} \tilde{a}_{22} & \tilde{a}_{23} \\ \tilde{a}_{32} & \tilde{a}_{33} \end{pmatrix}$,

as required.

The case of $x_{11} = \pm 1$ is considered by direct calculation.

4. Conclusions

The paper presents a proof of the hypothesis of J.-G. Sun, which is the basis of the theorem on the analytical properties of the eigenvalues of a matrix in the case of multiparameter perturbation of its elements.

5. Literature

1. Sun Ji-Guang. Eigenvalues and eigenvectors of a matrix dependent on several parameters. J. Comput. Math., 3, 351-364 (1985).
2. Bulkina E, Firsov A. Numerical-analytical method for solving the inverse problem of stability for technical systems with multiple uncertain parameters. "Machines. Technologies. Materials." International journal for science, technics and innovations for the industry (Bulgaria). Issue 11/2016, p. 12-14.
3. Kato T. A Short Introduction to Perturbation Theory for Linear Operators, N.-Y.: Springer-Verlag, 1982.
4. Beklemishev D.V. The course of analytical geometry and linear algebra, Moscow: FIZMATLIT, 2005 (in Russian).
5. Lancaster P., Tismenetsky M. The theory of matrices (2nd ed), N.-Y.: Academic Press, 1985.

THE COMPUTATION OF ROTOR'S MOTION IN CYLINDRICAL CHAMBER FILLED WITH VISCOUS GAS

РАСЧЕТ ДВИЖЕНИЯ РОТОРА В ЦИЛИНДРИЧЕСКОЙ КАМЕРЕ, ЗАПОЛНЕННОЙ ВЯЗКИМ ГАЗОМ

Prof., Dr. Dementev O.
Chelyabinsk State University, Chelyabinsk, Russia
e-mail: dement@csu.ru

Abstract: The problem rotor's movement in a stationary circular cylindrical chamber having finite length and filled with viscous gas is solved by the method of direct numerical integration of the set of equations describing pressure distribution in a thin layer of viscous gas and the motion of a rotating statically disbalanced cylinder. The rotor moving in the gravitational field is influenced by the impressed forces which vary periodically in time. Unsteady pressure equation is approximated by the symmetric stable finite-difference scheme of the second order accuracy. Stability criteria of a rotating rigid unstable cylinder (a rotor) motion subject to problem parameters are studied. The inner cylinder is influenced by outer forces which vary periodically in time. Trajectories of the rotor stationary motion for various velocities of rotation, disbalance values, amplitudes and frequencies of outer forces are calculated. Conditions of contact free motion of the cylinder, rotating in the chamber, are determined.

Keywords: Influence of form, stability, gas, rotation.

1. Let us consider two horizontal coaxial circular cylinders with length L and radius R_1 and R_2 . The space between the cylinders is filled with viscous gas. The center of mass of the inner rigid solid cylinder (rotor) situated outside its rotational axis (static instability). The rotor moves round its symmetry axis with constant angular speed ω . Outer cylinder (chamber) is immovable. A gap between the cylinders is significantly less than their radii that is why to determine pressure distribution in a thin layer of viscous gas one can use Reynolds equation. In cylindrical coordinate system (r, φ, z) , axis z of which is oriented along the axis of outer cylinder, the equation for pressure p is the following [1,2].

$$(1) \quad \frac{1}{12\mu} \frac{\partial}{\partial z} \left(h^3 \rho \frac{\partial p}{\partial z} \right) + \frac{1}{12\mu R_1^2} \left(h^3 \rho \frac{\partial p}{\partial \varphi} \right) = \frac{\partial}{\partial t} (\rho h) + \frac{\omega}{2} \frac{\partial}{\partial \varphi} (\rho h),$$

where ρ is gas density; $h = h(\varphi)$ – local thickness of a gap between cylindrical surfaces ($R_1 \leq r \leq R_1 + h$); μ – coefficient of dynamic viscosity of gas.

Boundary conditions:

$$(2) \quad \text{given } z = \pm L/2 \quad p = p_0$$

(p_0 – pressure in the medium around the layer). Having found the field of pressure in the gas layer from the problem (1), (2), let us find the force applied by the gas to the rotating cylinder of length L :

$$(3) \quad F_x = -2 \int_0^{L/2} \int_0^{2\pi} p R_1 \cos \varphi dz d\varphi, \\ F_y = -2 \int_0^{L/2} \int_0^{2\pi} p R_1 \sin \varphi dz d\varphi$$

While calculating reaction of a gas layer we took into account only pressure forces. They are much more than frictional forces [2] with the accuracy which was used while deriving the equation (1). Motion of a cylinder rotating in the gravitational field under the action of outer forces which periodically change in time is described by the equations (in the coordinate system, connected with the center of the immovable chamber).

$$(4) \quad m\ddot{x} = F_x + m\delta\omega^2 \cos \omega t + mg(1 + a_1 \cos \omega_1 t), \\ m\ddot{y} = F_y + m\delta\omega^2 \sin \omega t + mga_2 \sin \omega_1 t.$$

Here m – rotor mass; δ – value of shift of mass center from rotation axis; g – acceleration due to gravity; a_1, a_2 – amplitudes of outer

periodic effects (i.e., a case when metal cylinder moves in alternating electromagnetic field); ω_1 – frequency of outer effects. At the beginning rotation axis coincides with a chamber axis.

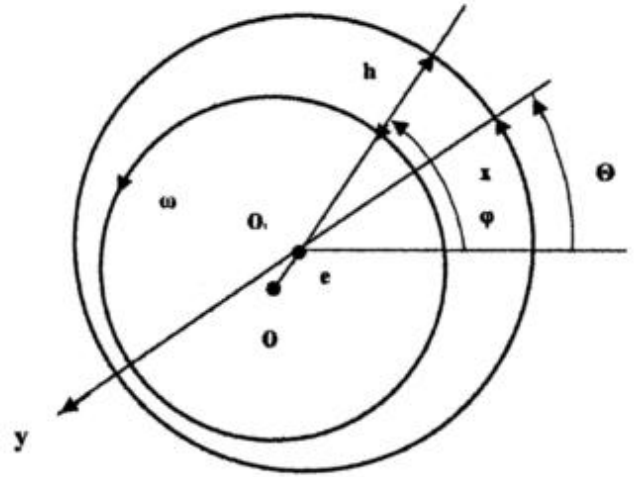


Fig. 1

2. Let us solve the problem of a rotor motion by the method of direct numerical integration of the set of equations describing a cylinder motion and pressure distribution in a gas layer [3, 4]. Let us rewrite the equation (1) in the form of the law of energy conservation of the volume $R_1 \Delta \varphi \Delta z h_j$; including the node i, j of the lattice in cylindrical coordinate system [5] ($\Delta \varphi$ – angular step dimension and Δz – is the step in a coordinate):

$$(5) \quad R_1 \Delta \varphi \Delta z \frac{\partial}{\partial t} (h_j \rho_{ij}) - \frac{h_j^3 \Delta \varphi R_1}{12\mu \Delta z} \left[\rho_{i+\frac{1}{2},j} (-p_{ij} + p_{i+1,j}) - \rho_{i-\frac{1}{2},j} (p_{ij} + p_{i-1,j}) \right] \left[\rho_{i,j+\frac{1}{2}} h_{j+\frac{1}{2}}^3 (p_{i,j+1} - p_{i,j}) - \rho_{i,j-\frac{1}{2}} h_{j-\frac{1}{2}}^3 (p_{ij} - p_{i,j-1}) \right] + \frac{\omega R_1 \Delta z}{2} \left[\rho_{i,j+\frac{1}{2}} h_{j+\frac{1}{2}} - \rho_{i,j-\frac{1}{2}} h_{j-\frac{1}{2}} \right] = 0 \\ (i = 0, 1, 2, \dots, I-1, \quad j = 1, 2, \dots, J).$$

here

$$\Delta\varphi = \frac{2\pi}{J}; \Delta z = \frac{L}{2I}; \varphi_j = j\Delta\varphi; h_j = C - x\cos\varphi_j - y\sin\varphi_j;$$

C – mean value of radial gap; I, J – quantity lattice nodes in axial and peripheral direction. Relation between pressure and gas density is defined by ratio $\frac{p}{\rho} = \text{const}$.

Boundary condition (2) in the form of finite difference are the following:

$$(6) \quad p_{i,j} = p_0, p_{-1,j} = p_{1,j}, p_{i,0} = p_{0,j}, p_{i,j+1} = p_{i,1} \\ (i = 1, 2, \dots, I, j = 1, 2, \dots, J).$$

Here $\Delta\varphi = 2\pi/J$; $\Delta z = L/2I$; $\varphi_j = j\Delta\varphi$; I, J – quantity of lattice nodes in axial and peripheral direction. Relation between pressure and gas density is defined by the ratio $p/\rho = \text{const}$.

While writing ratios we used conditions of periodicity

$$p(z, \varphi) = p(z, \varphi + 2\pi)$$

and smoothness $\left(\frac{\partial p}{\partial z}\right)_{z=0} = 0$.

Let us consider equations (5) in dimensionless form, using the following units of measurements: a distance across the layer – C (mean value of radial gap), a distance along the layer – R_1 , time – $1/\omega$, pressure – p_0 :

$$(7) \quad 2\Delta\varphi\Delta z \frac{\partial}{\partial t}(h_j \rho_{i,j}) - \\ h_j^3 \frac{\Delta\varphi}{\Delta z} [\rho_{i+1/2,j}(p_{i-1,j} - p_{i,j}) - \rho_{i-1/2,j}(p_{i,j} - p_{i+1,j})] \\ = \frac{\Delta z}{\Delta\varphi} [\rho_{i,j+1/2} h_{j+1/2}^3 (p_{i,j+1} - p_{i,j}) - \\ - h_{j-1/2}^3 \rho_{i,j-1/2} (p_{i,j} - p_{i,j-1})] + \\ + \Delta\Delta z (\rho_{i,j+1/2} h_{j+1/2}^3 - \rho_{i,j-1/2} h_{j-1/2}^3); \\ (h_j = 1 - x\cos\varphi_j - y\sin\varphi_j).$$

Using the method of varying direction [6] let us reduce the set problem of calculation of pressure field in a layer to a sequence of one-dimensional problems where we assign coefficients, non-linear terms and mass outlay of gas in axial direction to time moment $t_{k+1/2}$ (k – number of a time step) and other values – to t_k . Let us approximate the derivative $\partial p / \partial t$ by the difference formula of the second order centered with respect to $t_{k+1/4}$. We solve the derived nonlinear difference system of equations along the lattice lines considering index j as a parameter by the sweep method. Initial approach $p_{i,j}^0$ at every time step is defined with the help of linear extrapolation

$$(p_{i,j}^{k+1/2})^0 = 1, 5p_{i,j}^k - 0, 5p_{i,j}^{k-1} \quad k = 1, 2, \dots$$

where at the first time step $(p_{i,j}^{1/2})^0 = p_{i,j}^0$.

At the next stage of the method of varying directions in equations (7) let us refer linear components of mass outlays in peripheral directions to t_{k+1} and coefficients and other members of the equations to $t_{k+1/2}$, the derivative $\partial p / \partial t$ center with respect to

$t_{k+3/4}$. Then, following the method of the cyclic sweep method [7] find the pressure in all lattice nodes and the force effecting a rotor from a gas layer in the time moment t_{k+1} .

3. Let us consider the motion of a rotating unstable rotor in the field of gravity without periodic effects ($a_1 = a_2 = 0$). A rotor, being at the initial period of time in the center of a chamber, begins to move. Orbits of steady-state motion of the rotor are nearly circular ($\delta' = \delta/C = 0,3$; the velocity of rotation $n = 50 \text{ rev/s}$; $n = 350 \text{ rev/s}$).

Let us turn to studying motion of a rotor in the presence of set external periodic disturbances of the finite amplitude. Let us set disturbances only in vertical direction, i.e., we assume that $a_1 = 1$, $a_2 = 0$ and $\omega_1 = \omega$ (frequencies of the rotor motion and of external disturbances coincide). Calculations of trajectories of the rotor motion are carried out for the speed values of its rotation n in the interval from 50 to 3000 revolutions per second at the change of relative disbalance δ' from 0,1 to 1. Steady-state orbits of the rotor motion are close to vertical elliptic ones. It was found that at small ($n < n_1$) and big ($n > n_2$) velocities of rotation the motion of the cylinder was not stable; stationary closed trajectories of the axis of revolution were not formed and the rotor touched the chamber (Fig. 2). There is a limited zone of rotor velocities $[n_1, n_2]$ (Fig. 3) for a constant value of disbalance where it is possible for a rotor to move avoiding a contact with a chamber. Left boundary of a zone of steady-state orbits formation practically does not depend on relative disbalance and right boundary moves to the domain of big velocities of revolution with decreased disbalance. Domain ellipticity of the chamber (the function of a gap in this case had the form of (I8)) $h_j = 1,066 - 0,133\cos^2\varphi_j - x\cos\varphi_j - y\sin\varphi_j$ caused extension of formation zone for steady-state elliptic trajectories by means of increasing of the biggest critical velocity of revolution n_2 . When velocity increases by the factor of 10 (from 50 to 1500 rev/s) the maximal amplitude of the orbit a increases by the factor of 33.

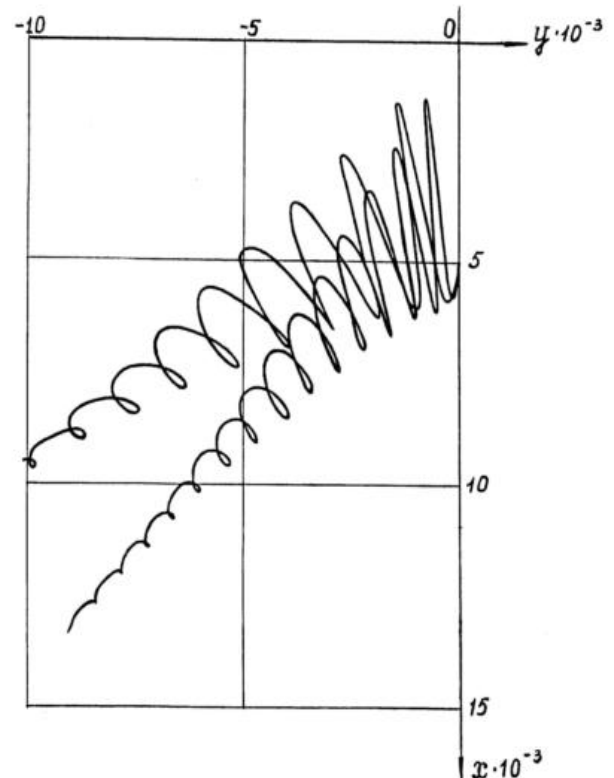


Fig. 2

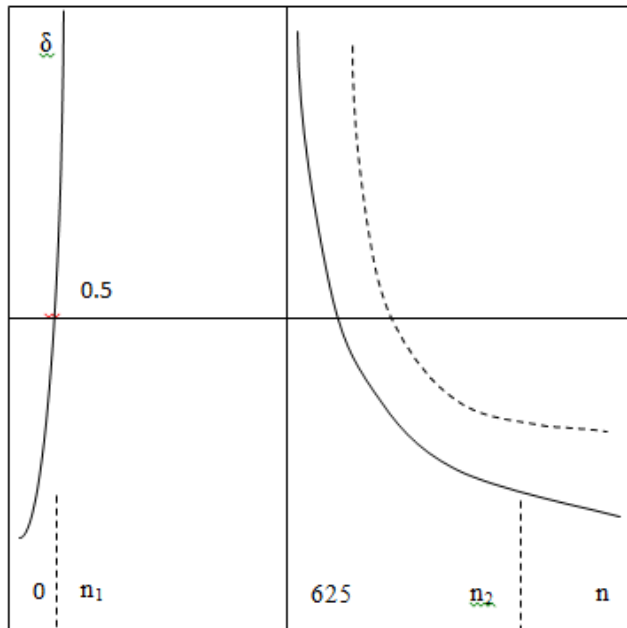


Fig.3

Increasing of the amplitude of horizontal disturbances a_2 up to 10 at low-frequency disturbances ($\omega_1 = 1s^{-1}$) actually does not distort steady-state circular trajectories of the rotor motion and does not cause their destruction. This conclusion confirms calculations carried out for velocities of cylinder rotation in the interval from 100 to 500 rev/s. At the amplitude of external effects equal to $a_2 = 100$

($a_1 = 0, \omega_1 = 1s^{-1}, n = 150$ rev/s) a cylinder gets in touch with a chamber very quickly ($\delta' = 0.3; a_1 = 0, a_2 = 1, \omega_1 = \omega = 2\pi 150 s^{-1}; a_1 = 0, a_2 = 0.1, \omega_1 = 1 s^{-1}, n = 150$ and 100 rev/s). Change of disbalance from 0,1 to 0,85 in the presence of vertical disturbances with the frequency equal a half of the rotor revolution frequency ($\omega = 200\pi s^{-1}, a_2 = 0, a_1 = 1$) does not influence significantly a formation and orientation of steady – state trajectories of the rotor. Influence of vertical disturbances of finite amplitude in comparison with the case of absence of disturbances causes increase of the amplitude of the steady-state trajectory by the factor of 8-10, shift of the orbit centre from the first quarter to the third quarter of the plane (x, y) and transformation of a circular orbit into an elliptic one.

$\omega = 200\pi s^{-1}, a_2 = 0, a_1 = 1$) does not influence significantly a formation and orientation of steady – state trajectories of the rotor. Influence of vertical disturbances of finite amplitude in comparison with the case of absence of disturbances causes increase of the amplitude of the steady-state trajectory by the factor of 8-10, shift of the orbit centre from the first quarter to the third quarter of the plane (x, y) and transformation of a circular orbit into an elliptic one.

REFERENCES

1. Kelzon A., Zimansky U., Yakovlev V. "The Dynamic of Rotor's in elastically Supports", Moscow, Nauka, 1982.
2. Konstantinesku V. "A Gaseouse bearings" [russian translation]. Mashinostroenie, Moscow, "MIR" 1968.
3. Tondl A. "Some Problems of Rotor Dynamics", Publishing House of the Czechoslovak Academy of Sciences, Prague, London, 1965.
4. Loitzansky L. "Mechanics of fluid and gas", Moscow, Nauka,, 1978.
5. Agishev G. "Research method of dynamic rotor's rotating in radial gaseous Bearings". Mashinovedenie. 1984; No 2. P. 131-157.
6. Samarsky A. "Introduction in theory of differences schemes", Moscow, Nauka, 1984.
7. Samarsky A., Nikolaev E. "Methods of solution of grid equations", Moscow, Nauka, 1978.
8. Dementiev O.N., Makhova V.A. "The computation of rotor's motion in cylindrical chamber". J. Appl. Mech. and Tech. Phys. (PMTF), No. 5. 1991, P. 68-72.

IMPLICIT EULER TIME DISCRETIZATION AND FDM WITH NEWTON METHOD IN NONLINEAR HEAT TRANSFER MODELING

Ph.D. Filipov S.¹, Prof. D.Sc. Faragó I.²

¹ Department of Computer Science, University of Chemical Technology and Metallurgy, Bulgaria

² Department of Applied Analysis and Computational Mathematics, MTA-ELTE Research Group, Eötvös Loránd University, Hungary
 filipovstefan@yahoo.com, faragois@cs.elte.hu

Abstract: This paper considers one-dimensional heat transfer in a media with temperature-dependent thermal conductivity. To model the transient behavior of the system, we solve numerically the one-dimensional unsteady heat conduction equation with certain initial and boundary conditions. Contrary to the traditional approach, when the equation is first discretized in space and then in time, we first discretize the equation in time, whereby a sequence of nonlinear two-point boundary value problems is obtained. To carry out the time-discretization, we use the implicit Euler scheme. The second spatial derivative of the temperature is a nonlinear function of the temperature and the temperature gradient. We derive expressions for the partial derivatives of this nonlinear function. They are needed for the implementation of the Newton method. Then, we apply the finite difference method and solve the obtained nonlinear systems by Newton method. The approach is tested on real physical data for the dependence of the thermal conductivity on temperature in semiconductors. A MATLAB code is presented.

Keywords: HEAT CONDUCTION EQUATION, TEMPERATURE-DEPENDENT THERMAL CONDUCTIVITY, IMPLICIT EULER METHOD, BOUNDARY VALUE PROBLEM, FINITE DIFFERENCE METHOD, NEWTON METHOD

1. Introduction

We consider the one-dimensional unsteady heat conduction equation [1-3]

$$\rho c_p \frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(\kappa(u) \frac{\partial u}{\partial x} \right), \quad (1)$$

where $u(x, t)$ is the temperature at position x and time t , ρ is the density, c_p is the heat capacity at constant pressure, and κ is the thermal conductivity of the media. We assume that ρ and c_p have constant values, but κ depends on the temperature u . This assumption is often justifiable for certain temperature range (e.g. for silicon [4]). Performing the differentiation in the right-hand side of (1) we get

$$\rho c_p \frac{\partial u}{\partial t} = \partial_u \kappa(u) \left(\frac{\partial u}{\partial x} \right)^2 + \kappa(u) \frac{\partial^2 u}{\partial x^2}. \quad (2)$$

When κ does not depend on u , i.e. $\partial_u \kappa(u) = 0$, then (2) is a linear (parabolic) partial differential equation. When $\partial_u \kappa(u) \neq 0$, then (2) is nonlinear. Equation (2) will be solved on the spatial interval $[a, b]$ subject to certain boundary and initial conditions:

$$u(a, t) = \alpha(t), \quad u(b, t) = \beta(t), \quad t > 0, \quad (3)$$

$$u(x, 0) = u_0(x), \quad x \in [a, b]. \quad (4)$$

The boundary conditions (3) give the temperature at the two ends as function of time. The initial condition (4) specifies the initial spatial distribution of the temperature.

2. Implicit Euler time discretization

For the linear problem ($\partial_u \kappa(u) = 0$), the numerical methods for solving (2)-(4) are well elaborated (finite difference, finite element methods, etc.). Usually, (2) is first discretized in space, whereby an initial-value (Cauchy) problem for first order ODE system is obtained. If the explicit Euler method is used to solve the Cauchy problem, then the method is stable only for $0 < \Delta\tau/h^2 \leq$

$1/2$, where $D = \kappa/(\rho c_p)$ is the thermal diffusivity, h is the discretization step in space, and τ is the discretization step in time.

Our approach is different. We first discretize (2) in time. Using a time-step τ , the time line $t \geq 0$ is partition by equally separated mesh-points:

$$t_n = n\tau, \quad n = 0, 1, 2, \dots \quad (5)$$

Then, using implicit Euler scheme [5], equation (2) is discretized on the mesh (5):

$$\rho c_p \frac{u_n - u_{n-1}}{\tau} = \partial_u \kappa(u_n) \left(\frac{du_n}{dx} \right)^2 + \kappa(u_n) \frac{d^2 u_n}{dx^2}, \quad (6)$$

where $u_n = u_n(x)$ and $u_{n-1} = u_{n-1}(x)$ approximate the values of $u(x, t_n)$ and $u(x, t_{n-1})$, respectively. Equation (6) approximates the partial differential equation (2). The error is $O(\tau)$, hence the discretization scheme is first-order accurate in time. The method is stable, unlike the explicit method, i.e. evaluating the right-hand side of (2) at u_{n-1} , which is only conditionally stable.

Solving (6) for the second spatial derivative of the temperature u_n we get

$$\frac{d^2 u_n}{dx^2} = f(u_n, v_n; u_{n-1}), \quad (7)$$

where $v_n = du_n/dx$ is the temperature gradient and f is the following nonlinear function:

$$f(u_n, v_n; u_{n-1}) = \frac{\phi(u_n, v_n; u_{n-1})}{\kappa(u_n)}, \quad (8)$$

$$\phi(u_n, v_n; u_{n-1}) = \rho c_p \frac{u_n - u_{n-1}}{\tau} - \partial_u \kappa(u_n) v_n^2. \quad (9)$$

Equation (7), together with the boundary conditions $u_n(a) = \alpha(t_n)$, $u_n(b) = \beta(t_n)$, constitutes a nonlinear two-point boundary value problem (TPBVP) for the unknown function u_n . If the function u_{n-1} is known (given) the problem can be solved by using some numerical technique for nonlinear problems. Thus, starting

from the initial condition u_0 we can solve successively (7) for $n = 1, 2, \dots$

3. Derivatives for the Newton method

The implementation of the Newton method requires the partial derivatives of $f(u_n, v_n; u_{n-1})$ with respect to u_n and v_n . Introducing the notation $f_n = f(u_n, v_n; u_{n-1})$, $\phi_n = \phi(u_n, v_n; u_{n-1})$ and denoting the derivatives by $q_n = q(u_n, v_n; u_{n-1})$, $p_n = p(u_n, v_n)$ we get:

$$q_n = \frac{\partial f_n}{\partial u_n} = \frac{1}{\kappa(u_n)} \left(-f_n \partial_u \kappa(u_n) + \frac{\partial \phi_n}{\partial u_n} \right), \quad (10)$$

$$p_n = \frac{\partial f_n}{\partial v_n} = \frac{1}{\kappa(u_n)} \frac{\partial \phi_n}{\partial v_n}, \quad (11)$$

where

$$\frac{\partial \phi_n}{\partial u_n} = \frac{\rho c_p}{\tau} - \partial_{uu}^2 \kappa(u_n) v_n^2, \quad (12)$$

$$\frac{\partial \phi_n}{\partial v_n} = -2 \partial_u \kappa(u_n) v_n. \quad (13)$$

4. Finite difference method

Since $\rho c_p / \tau$ grows to infinity as τ goes to zero (which effects IVP solutions), it turns out that the finite difference method (FDM) [6] is a better choice for the solution of the obtained TPBVPs than the shooting method [6,7]. Hence, we adopt the FDM. The interval $[a, b]$ is partitioned by N equally separated mesh-points:

$$x_i = a + (i-1)h, i = 1, 2, \dots, N, h = \frac{b-a}{N-1}. \quad (14)$$

Equation (7) is discretizes on the uniform mesh (14) using the FDM with the central difference approximation:

$$\frac{u_{n,i+1} - 2u_{n,i} + u_{n,i-1}}{h^2} = f(u_{n,i}, v_{n,i}; u_{n-1,i}), \quad (15)$$

$$i = 2, 3, \dots, N-1. \quad (16)$$

Correspondingly, everywhere in equations (8)-(13), we set $x = x_i$, and then replace the values of $u_n(x_i)$, $v_n(x_i)$, and $u_{n-1}(x_i)$ with their approximations $u_{n,i}$, $v_{n,i}$, and $u_{n-1,i}$ where

$$v_{n,i} = \frac{u_{n,i+1} - u_{n,i-1}}{2h}. \quad (17)$$

Equation (15) approximates (7) with error $O(h^2)$, i.e. it is second-order accurate in space. Equation (15) holds for the inner mesh-points. At the boundaries we apply the boundary conditions:

$$u_{n,1} = \alpha(t_n), u_{n,N} = \beta(t_n) \quad (18)$$

5. Solving the nonlinear system by Newton method

Introducing the column-vector $\mathbf{G}_n = [G_{n,1}, G_{n,2}, \dots, G_{n,N}]^T$ with components

$$G_{n,1} = u_{n,1} - u_a(t_n), G_{n,N} = u_{n,N} - u_b(t_n), \quad (19)$$

$$G_{n,i} = u_{n,i+1} - 2u_{n,i} + u_{n,i-1} - h^2 f_{n,i}, \quad (20)$$

$$f_{n,i} = f(u_{n,i}, v_{n,i}; u_{n-1,i}), \quad (21)$$

the system of nonlinear equations (15) and the boundary conditions (18) are written as one equation:

$$\mathbf{G}_n(\mathbf{u}_n) = 0, \quad (22)$$

where

$$\mathbf{u}_n = [u_{n,1}, u_{n,2}, \dots, u_{n,N}]^T. \quad (23)$$

Starting by some initial guess $\mathbf{u}_n^{(0)}$, the nonlinear system (22) can be solved by the Newton iterative method:

$$\mathbf{u}_n^{(k+1)} = \mathbf{u}_n^{(k)} - (\mathbf{L}_n^{(k)})^{-1} \mathbf{G}_n(\mathbf{u}_n^{(k)}),$$

$$k = 0, 1, 2, \dots \quad (24)$$

where $\mathbf{L}_n^{(k)}$ is the Jacobian of $\mathbf{G}_n$ with respect to $\mathbf{u}_n$ evaluated at $\mathbf{u}_n^{(k)}$:

$$\mathbf{L}_n^{(k)} = \frac{\partial \mathbf{G}_n}{\partial \mathbf{u}_n}(\mathbf{u}_n^{(k)}). \quad (25)$$

Calculating the elements of the Jacobian we get:

$$L_{n(1,1)}^{(k)} = 1, \quad L_{n(N,N)}^{(k)} = 1, \quad (26)$$

$$L_{n(i,i)}^{(k)} = -2 - h^2 q_{n,i}^{(k)}, \quad L_{n(i,i-1)}^{(k)} = 1 + \frac{1}{2} h p_{n,i}^{(k)},$$

$$L_{n(i,i+1)}^{(k)} = 1 - \frac{1}{2} h p_{n,i}^{(k)}, \quad (27)$$

$$q_{n,i}^{(k)} = q(u_{n,i}^{(k)}, v_{n,i}^{(k)}; u_{n-1,i}^{(k)}), p_{n,i}^{(k)} = p(u_{n,i}^{(k)}, v_{n,i}^{(k)}). \quad (28)$$

Iteration (24) is a one-step (two-level) iteration. Starting from some initial guess $\mathbf{u}_n^{(0)}$, we can find each next approximation $\mathbf{u}_n^{(k+1)}$, $k = 0, 1, 2, \dots$ using (24). If the sequence is convergent, then the limiting vector $\mathbf{u}_n = \lim_{k \rightarrow \infty} (\mathbf{u}_n^{(k+1)})$ is a solution to the nonlinear system (22). In practice, the iteration process is usually ended when

$$\|\mathbf{u}_n^{(k+1)} - \mathbf{u}_n^{(k)}\| < \epsilon. \quad (29)$$

This inequality is called a stopping criteria. The vector $\mathbf{u}_n^{(k+1)}$ is taken as approximate solution to (22). As an initial guess $\mathbf{u}_n^{(0)}$, we can use the solution $\mathbf{u}_{n-1}$ found at the previous step.

6. Computer experiment

Consider a thin homogenous rod, along the x -axis between the points $x = 1$ and $x = 3$, without heat sources and without radiation. The density ρ and the heat capacity c_p are constant, but the thermal conductivity κ depends on the temperature as

$$\kappa = \kappa_0 \exp(\chi u). \quad (30)$$

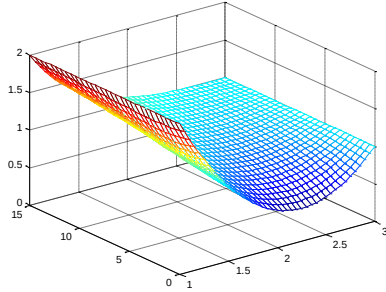
Such a temperature dependence actually occurs in real physical systems, e.g. for silicon [4]. We choose the following values of the constants: $\rho = 1$, $c_p = 1$, $\kappa_0 = 0.1$. The temperature at the two ends is kept constant:

$$u(1, t) = 2, u(3, t) = 1, t > 0. \quad (31)$$

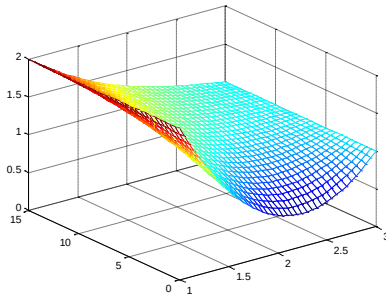
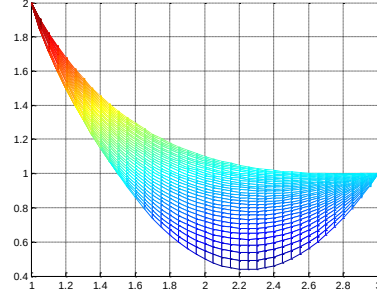
The initial temperature profile is

$$u(x, 0) = 2 - \frac{x-1}{2} + (x-1)(x-3), x \in [1, 3]. \quad (32)$$

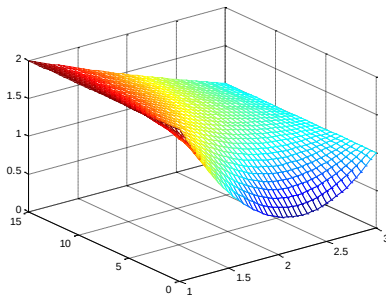
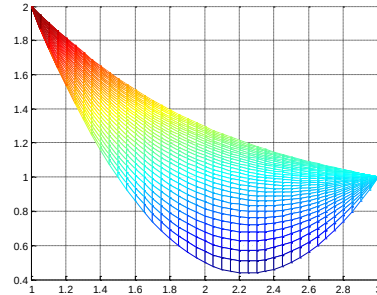
To find the time evolution of (32), we solve the partial differential equation (1) with boundary conditions (31) and initial conditions (32) by the method described in this paper. The equation is solved for $\chi = -1.0, -0.5, 0, 0.5, 1.0$. The step-size is chosen to be $\tau = 0.5$ with integration range $0 \leq t \leq 15$. The spatial interval $x \in [1, 3]$ is discretized by $N = 41$ mesh-points, i.e. $h = 0.05$. The results are shown in Fig. 1



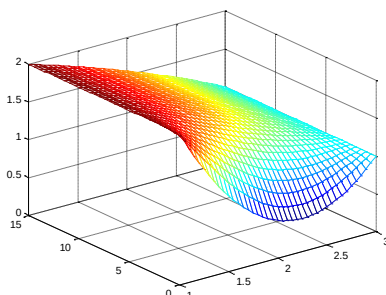
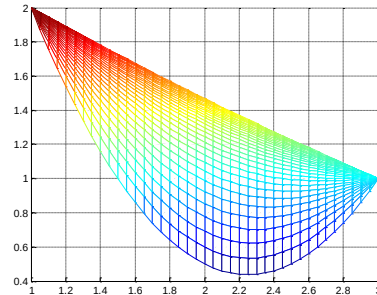
$\chi = -1.0$



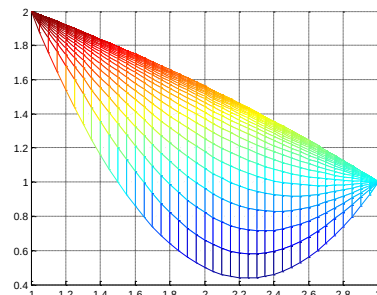
$\chi = -0.5$



$\chi = 0$



$\chi = 0.5$



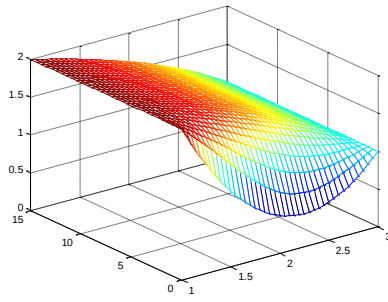
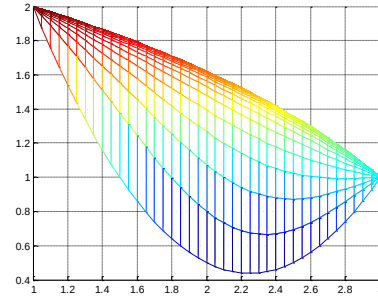

 $\chi = 1.0$


Fig. 1. Solving the heat conduction equation (1) for boundary conditions (31), initial condition (32), and thermal conductivity (30). In addition to the 3D view (left), a side-view u vs. x is also shown (right).

For $\chi = 0, 0.5$, and 1.0 the final temperature distribution reached in the experiment is practically the steady-state distribution. For $\chi = -1.0$ and -0.5 a little bit more time is needed. The steady-state distribution for $\chi = 0$ is, as expected, linear.

7. Conclusion

This paper considered heat transfer with temperature-dependent thermal conductivity. The one-dimensional unsteady heat conduction equation was solved numerically by using implicit time-discretization and FDM with Newton method for the solution of the arising nonlinear two-point boundary value problems. Data for the dependence of the thermal conductivity on temperature in certain semiconductors was used. The results obtained by the numerical computer experiments are consistent with the expected outcome. The proposed method is stable, unlike its explicit counterpart.

8. Appendix

A MATLAB code is presented for the numerical solution of the example provided in section 6.

```
function main
rho=1; Cp=1; kappa0=0.1; chi=0.5;
M=31; N=41;
tEnd=15; tau=tEnd/(M-1);
a=1; b=3; h=(b-a)/(N-1);
alpha=2; beta=1;
x=zeros(N,1);
u0=zeros(N,1);
for i=1:N
    x(i)=a+(i-1)*h;
    u0(i)=2-(x(i)-1)/2+(x(i)-1)*(x(i)-3);
end
t=zeros(M,1);
for n=1:M
    t(n)=(n-1)*tau;
end
u_1=zeros(N,1); u=zeros(N,1);
uNext=zeros(N,1); G=zeros(N,1);
L=zeros(N,N); L(1,1)=1; L(N,N)=1;
U=zeros(N,M); U(:,1)=u0;
for n=2:M
    u=U(:,n-1); u_1=U(:,n-1);
    eps=1;
    while(eps>0.0001)
        G(1)=u(1)-alpha;
        G(N)=u(N)-beta;
        for i=2:N-1
            k=kappa0*exp(chi*u(i));
            Dk=chi*k;
            D2k=chi*Dk;
            v=(u(i+1)-u(i-1))/(2*h);
            A=rho*Cp/tau;
            phi=A*(u(i)-u_1(i))-Dk*v*v;
            f=phi/k;
            q=(-f*Dk+A-D2k*v*v)/k;
```

```
p=-2*Dk*v/k;
G(i)=u(i+1)-2*u(i)+u(i-1)-h*h*f;
L(i,i-1)=1+0.5*h*p;
L(i,i)=-2-h*h*q;
L(i,i+1)=1-0.5*h*p;
end
uNext=u-L\G;
eps=sqrt(h*(uNext-u)*(uNext-u));
u=uNext;
end
U(:,n)=u;
end
mesh(x,t,U);
end
```

The mathematical quantities and the corresponding variables used in the MATLAB code are shown in Table 1.

Paper	MATLAB
ρ	rho
c_p	Cp
κ_0	kappa0
χ	chi
τ	tau
N	N
a	a
b	b
h	h
α	alpha
β	beta
x_i	x(i)
t_n	t(n+1)
u_0	u0
$u_n^{(k)}$	u
$u_n^{(k+1)}$	uNext
u_{n-1}	u_1
$u_{n,i}$	U(i,n+1)
$G_n(u_n^{(k)})$	G
$L_n^{(k)}$	L
$\ u_n^{(k+1)} - u_n^{(k)}\ $	eps
$u_n^{(k)}$	u(i)
$\kappa(u_{n,i}^{(k)})$	k
$\partial_u \kappa(u_{n,i}^{(k)})$	Dk
$\partial_{uu}^2 \kappa(u_{n,i}^{(k)})$	D2k
$v_{n,i}^{(k)}$	v
$\phi(u_{n,i}^{(k)}, v_{n,i}^{(k)}, u_{n-1,i})$	phi
$f(u_{n,i}^{(k)}, v_{n,i}^{(k)}, u_{n-1,i})$	f
$q_{n,i}^{(k)}$	q
$p_{n,i}^{(k)}$	p

Table 1.

9. Acknowledgement

This research was carried out in the ELTE Institutional Excellence Program (1783-3/2018/FEKUTSRAT) supported by the Hungarian Ministry of Human Capacities, and supported by the Hungarian Scientific Research Fund OTKA, No. K112157 and SNN125119

10. References

- [1] R.B. Bird, W.E. Stewart, E.N. Lightfoot, *Transport Phenomena*, 2nd Edition, John Wiley & Sons, Inc., (2002)
- [2] H.S. Carslaw, J.C. Jaeger, *Conduction of Heat in Solids*, 2nd Edition, Oxford University Press, (1986)
- [3] S.L. Sobolev, *Partial Differential Equations of Mathematical Physics*, Dover Publications, (1989)
- [4] J. Lienemann, A. Yousefi, J.G. Korvink, *Nonlinear Heat Transfer Modeling*, In: P. Benner, D.C. Sorensen, V. Mehrmann (eds) *Dimension Reduction of Large-Scale Systems*, Lecture Notes in Computational Science and Engineering, 45 (2005) 327-331
- [5] U.M. Ascher, S.J. Ruuth, R.J. Spiteri: *Implicit-Explicit Runge-Kutta Methods for Time-Dependent Partial Differential Equations*, *Appl Numer Math*, vol. 25 (2-3) (1997)
- [6] U.M. Ascher, R.M.M. Mattheij, R.D. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, in: *Classics in Applied Mathematics*, vol. 13, SIAM, (1995)
- [7] S.M. Filipov, I.D. Gospodinov, I. Faragó, *Shooting-projection method for two-point boundary value problems*, *Appl. Math. Lett.* 72 (2017) 10-15
<https://doi.org/10.1016/j.aml.2017.04.002>

MATHEMATICAL MODELS OF CALCULATIONS OF PARAMETERS OF CRYSTALLIZATION OF BINARY ALLOYS BY MEANS OF COMPUTER THERMAL ANALYSIS

Ass. Prof., Dr. Eng. Donii O.¹, Ass. Prof., Dr. Eng. Khristenko V.¹, Omelko L.², Ass. Kotliar S.¹

National Technical University of Ukraine "Kyiv Polytechnic Institute named after Igor Sikorsky" - Kyiv, Ukraine¹
Physical Technological Institute of Metals and Alloys, National Academy of Science of Ukraine²

Email: dosha@iff.kpi.ua

Abstract: Mathematical models of crystallization of binary metal alloys crystallizing in the temperature range are proposed. These models are intended for use in the composition of quality control systems for cast alloys based on computer thermal analysis. The proposed models make it possible to calculate the time dependences of the relative amounts of the solid and liquid phases using the data of the cooling curve, as well as to determine the intensity of the increase of the amount of the solid phase. The method of determining the temperature dependences of the specific heat and latent crystallization heat using the data on the temperature-concentration dependences of the free energies of the phases forming the system under study was used.

KEYWORDS: THERMAL ANALYSIS, CRYSTALLIZATION, ENTHALPHY, GIBBS ENERGY, HEAT CAPACITY

1. Introduction

The final level of service properties of cast parts is laid down during the smelting and crystallization of alloys. It is at the stage of crystallization the "primary" structure of the solid metal is formed, which essentially determines the properties of the final product. Therefore, the creation of systems for operational control and management of the state of alloy, as well as control and management of crystallization phase of the melt is relevant.

The operation of many systems for monitoring of the state of metal melts is based on thermal analysis, which provides information on the thermal effects accompanying cooling of sample [1–7]. On the cooling curves (dependences of temperature on time $T = f(t)$), the thermal effects accompanying the formation of the cast metal structure are especially pronounced. Therefore, it seems promising studying the kinetics of crystallization and prediction of the service properties of metals in the solid state from the results of the analysis of cooling curves. The information which is obtained by this way (especially in the express mode) is necessary for making operational decisions on management of casting technological processes [5, 6].

In [8, 9], mathematical models were developed that describe the solidification of pure metals based on computer thermal analysis in the subsystem of quality control of casting melts. They describe the crystallization of a small sample of metal (for sufficiently small values of the Biot criterion: $Bi \ll 1$). When they were created, it was assumed that the specific heat (c) and latent heat of crystallization (L) are constants. Indeed, during the crystallization of pure metals, the specific latent heat of crystallization L is a constant value, while specific heat capacities of the liquid and solid phases differ (for example, for aluminum - by about 20%). It should be noted that these models were also used to describe the crystallization of alloys [8 - 10]. However, the crystallization of alloys occurs in the temperature range with changes in the composition of both the liquid and solid phases. Therefore, the values of the specific heat c and the specific latent heat of crystallization L depend on the composition and temperature of the liquid-solid medium, which, in turn, changes with time. Therefore, in the mathematical models of crystallization, intended for use in thermal analysis, it is necessary to make appropriate corrections.

The purpose of this work is to develop mathematical models of crystallization of multicomponent metal melts, which are crystallized in the temperature range, and are intended to interpret the results of computer thermal analysis.

2. Solution of the problem under consideration

Considering the solidification of a small casting, when the condition $Bi \ll 1$ is satisfied (which is typical for thermal analysis), we can write the heat balance equation [8]:

$$dQ = dQ_c + dQ_L, \quad (1)$$

where dQ is the amount of heat given by the alloy to the

environment; dQ_L - the amount of heat released as a result of the

formation of a solid phase; dQ_c - the amount of heat released as a result of temperature changes during the cooling of the metal.

Such a record of the heat balance equation (1) assumes that the temperature gradient over the cross section of the casting is neglected, and the casting is a homogeneous material point with mass m_0 .

In [6, 8 - 10], the well-known calorimetric relations were substituted in (1):

$$dQ_c = -cm_0dT(t), \quad (2)$$

$$dQ_L = Ldm(t), \quad (3)$$

and the intensity (speed) of heat removal to the environment was described by the law of total heat transfer:

$$\frac{dQ}{dt} = fS[T(t) - T_e] + \sigma\epsilon S[T^4(t) - T_e^4], \quad (4)$$

where $T(t)$ is the function of the dependence of the temperature of the metal sample on time (cooling curve of thermal analysis); T_e is the ambient temperature; c is the specific heat, L is the specific latent heat of crystallization; m_0 is the mass of the metal under study; $m(t)$ is the mass of the metal which have been crystallized to the moment of time t ; f is the heat transfer coefficient; S is the surface of the area from which heat is removed to the environment; ϵ - is the degree of blackness of the sample's surface; σ is the Stefan - Boltzmann constant.

Taking into account that the thermophysical parameters of alloys in the crystallization process are changed, it seems appropriate to determine the temperature-concentration dependences of the heat capacities of the phases $c(x, T(t))$ and the specific latent heat of crystallization $L(x, T(t))$ in the temperature range of crystallization for the particular alloy under study (here x is the content of the second component in the melt).

In [11], a method was proposed for determining of the dependences of interest from analytical expressions for the temperature-concentration dependences of the molar free energies of the phases. The calculation is based on the assumption that the values of enthalpy, entropy and isobaric-isothermal potential are interconnected through the heat capacity C_p :

$$\left. \begin{aligned} H &= \int c_p dT, \\ S &= \int \frac{c_p}{T} dT, \\ G &= H - TS = \int c_p dT - T \int \frac{c_p}{T} dT, \end{aligned} \right\} \quad (5)$$

where G is the Gibbs free energy; H - is enthalpy; S - entropy. For many metallic systems data on the temperature-concentration dependences of the Gibbs free energy of the phases forming the system under consideration are given in the literature. For example, for the case of a two-component system, the temperature-concentration dependence of the Gibbs molar energy of the phase is defined as:

$$G(x, T) = G_A(T) \cdot (1 - x_B) + G_B(T) \cdot x_B + G_{mix}(x, T) + RT \cdot (x_A \ln x_A + x_B \ln x_B), \quad (6)$$

where $G_A(T)$, $G_B(T)$ are the molar free energies of the pure components A and B in the phase at temperature T ; $x_A = 1 - x_B$, x_B - are content of components A and B in the phase (molar fractions); $G_{mix}(x, T)$ - is excess molar free energy of mixing in the phase of composition x at temperature T ; R - is the universal gas constant.

If the temperature-concentration dependence of the molar heat capacity $c_p(x, T)$ of the phase is expressed as a polynomial

$$c_p(x, T) = z_3(x) + z_4(x)T + \frac{z_5(x)}{T^2} + z_6(x)T^2 + z_7(x)T^3 + \dots, \quad (7)$$

then, in accordance with (5):

$$G(x, T) = z_1 - z_2T + z_3(x)T(1 - \ln T) - \frac{z_4(x)T^2}{2} - \frac{z_5(x)}{2T} - \frac{z_6(x)T^3}{6} - \frac{z_7(x)T^4}{12} + \dots, \quad (8)$$

where z_1 and z_2 are the integration constants.

The dependence of the heat capacity on the composition can be taken into account by presenting the coefficients in (7) in the form of polynomials

$$z_k(x) = z_{k,1}(x) + z_{k,2}(x)x + z_{k,3}(x)x^2 + \dots \quad (9)$$

Therefore, by presenting the temperature-concentration dependences of the free energies of the phases (6) in the form of polynomials (8) (for example, using the polynomial regression), one can obtain the desired temperature-concentration dependences of the heat capacities of the phases $c(x, T(t))$.

Taking into account (5), the temperature-concentration dependences of the enthalpies of the phases take the form:

$$H(x, T) = z_1 + z_3(x)T + \frac{z_4(x)T^2}{2} - \frac{z_5(x)}{T} + \frac{z_6(x)T^3}{3} + \frac{z_7(x)T^4}{4} + \dots, \quad (10)$$

Changes of the enthalpy of the system during the transition from the melt to the solid solution, calculated as the difference between the enthalpies of the liquid and solid phases (taken with the appropriate sign), are nothing else but the desired temperature-concentration dependences of the specific heat of crystallization $L(x, T(t))$.

Having determined by this way the temperature-concentration dependences $c(T(t))$ and $L(T(t))$ and substituting them in (2) and (3), equation (1) can be reduced to the form:

$$\frac{dV(t)}{dt} = \frac{c(T(t))}{L(T(t))} \frac{dT(t)}{dt} + \frac{k_{co}}{L(T(t))} [T(t) - T_e] + \frac{k_{ra}}{L(T(t))} [T^4(t) - T_e^4], \quad (11)$$

$$V(t) = \frac{m(t)}{m_0}, \quad (12)$$

$$k_{co} = \frac{fS}{m_0}, \quad (13)$$

$$k_{ra} = \frac{\sigma \mathcal{E}}{m_0}, \quad (14)$$

where $V(t)$ is the relative amount of solids.

The solution of equation (11) can be obtained by direct integration:

$$V(t) = \int_w^t \frac{c(T(t))}{L(T(t))} \frac{dT(t)}{dt} dt + \int_w^t \frac{k_{co}}{L(T(t))} [T(t) - T_e] dt + \int_w^t \frac{k_{ra}}{L(T(t))} [T^4(t) - T_e^4] dt. \quad (15)$$

The under integral functions in (15) are complex non-linear dependencies. However, taking into account the fact, that the result of thermal analysis is the cooling curve $T(t)$, which is a time series with constant time intervals [9], the integrals in expression (15) are easily calculated by numerical methods. The calculation uses the initial condition:

$$V(t_{cr}) = 0, \quad (16)$$

showing the absence of a solid phase at the onset of crystallization.

3. Results and discussion

The proposed mathematical model (16) allows to determine the amount of the solid phase $V(t)$ at any time of crystallization using the cooling curve obtained by thermal analysis. Although the same information can be obtained from the alloy state diagram, due to the non-equilibrium crystallization conditions, its use leads to significant calculation inaccuracies. Beside this, the usage of a state diagram does not allow to estimate the intensity (rate) of the formation of a solid phase $V'(t)$. But this is namely that parameter which is used for prediction of mechanical properties of alloys [5, 6].

For practical calculations of the amount of the solid phase and the intensity of its formation using expressions (11) and (16), first of all it is necessary to determine the time derivative of the temperature $T'(t)$ from the cooling curve and calculate the coefficients k_{co} and k_{ra} . Calculation of $T'(t)$ is easily implemented using the numerical differentiation formulas. However, in many cases, pre-filtering of the cooling curve is required [9]. To determine the values of the coefficients k_{co} and k_{ra} from the cooling curve, one can use the method of determining of the coefficients k_1 and k_2 , given in [9]. In [9], they are calculated by linearizing the portions of the cooling curve before and after crystallization, followed by the using of the least squares method. The coefficients k_{co} and k_{ra} are related to k_1 and k_2 as follows:

$$k_{co} = k_1 c_{Li}, \quad (11)$$

$$k_{ra} = k_1 c_{So}, \quad (12)$$

where c_{Li} , c_{So} are the specific heat values of the studied alloy in the liquid and solid states.

The above mentioned methodology can also be used in the development of mathematical models of crystallization of alloys with eutectic. Models for such alloys are given in [10], where temperature dependences of thermal parameters are not taken into account.

4. Conclusions

In this paper, we developed a method for constructing of

mathematical models of crystallization of binary metal alloys which are crystallized in the temperature range. A characteristic feature of the proposed models is to take into account the dependence of the specific heat and latent heat of crystallization on temperature. For this purpose, the temperature-concentration dependences of the free energies of the phases forming the system under study are used. The proposed models make it possible to calculate the time dependences of the relative amounts of the solid and liquid phases from the cooling curves obtained by using of computer thermal analysis, as well as to determine the intensity of formation of the solid phase.

5. References

1. Apelian D., Sigworth G.K., Whaler K.R. Assessment of grain refinement and modification of Al-Si foundry alloys by thermal analysis // Trans. Amer. Foundrymen's Soc. - 1984. - 92. - p. 297-307.
2. Argyropoulos S., Closset B., Gruzelski J.E., Oger H. Quantitative Kontrolle der verdellung von Al-Si gublegierungen unter anwendung eine thermoanalysetechnik // Giesser. Prax. - 1985. No11. - s. 173-180.
3. Argyropoulos S., Closset B., Gruzelski J.E., Oger H. The quantitative control of modification in Al-Si foundry alloys using thermal analysis technique // Transact. Amer. Foundrymen's Soc. - 1983. - 91. - p. 351-358.
4. Shumikhin V.S., Vitusevich V.T., Kornienko G.L. Comprehensive quality control of cast iron by thermal analysis // Foundry. - 1984. - No2. - P. 3-5 (Russian)
5. Bialik OM, Donii A.N., Pikovsky V.S., Shapoval A.I. The usage of the computer system for quality control of melts - "Foundry", 1996, No. 2, pp.28-30. (Russian)
6. Bialik OM, Donii A.N., Shapoval A.I., Kulinich A.A. System of computer thermoanalysis of silumin - "Foundry production", 1998, №7, p.41-42. (Russian)
7. Borisov G. P., Smulsky A. A., Semenchenko A. I. Express-control of the melt and prediction of the properties of a future casting at the stage of preparation of a liquid metal based on an improved method of thermal analysis // Casting processes. - 2007. - № 1-2. - p. 19-22. (Russian)
8. Bialik OM, Mentkovsky Yu.L. Questions of the dynamic theory of solidification of metal castings. - Kiev: Vishcha school, 1983. - 111 p. (Russian)
9. Donii O.M. Mathematical models for calculating of the parameters of crystallization and preliminary filtration of the cooling curve during computer thermal analysis // Bulletin of SevNTU. Vip 110: Mechanics, energy, ecology: zb. sciences. pr. Sevastop. nat tech. un-t - Sevastopol: View of SevNTU, 2010. - p. 193-197. (Ukrainian)
10. Donii O.M. . Mathematical models of hardening of alloys with eutectic for calculation of parameters of crystallization based on data of the cooling curve in computer thermal analysis // Bulletin of SevNTU. Vip 110: Mechanics, energy, ecology: zb. sciences. pr. Sevastop. nat tech. un-t - Sevastopol: View of SevNTU, 2011. - p. 126 - 131. (Ukrainian)
11. Khristenko V., Omelko L., Donii O. Calculation of temperaure dependence of thermal effects at cooling of metal alloys // Machines. Technologies. Materials. - 9. - 2018. - P.374 - 377.

MODELING OF MAGNETIC HYSTERESIS UNDER WEAK MAGNETIC FIELDS AND TRIAXIAL STRESS STATE

Mushnikov A.N., Kryucheva K. D.

Institute of Engineering Science, Ural Branch of the Russian Academy of Sciences, Ekaterinburg, Russian Federation.

mushnikov@imach.uran.ru

Abstract: The magnetomechanical hysteresis models of Jiles-Atherton and its modifications by Sablik are extended to treat magnetic properties in the case of triaxial stress state. Unlike the previous version, it is focused on weak magnetic fields. Results of simulation are compared with experimental data.

Keywords: MAGNETIC HYSTERESIS, BULK STRESS STATE

1. Introduction

The problem of a theoretical description of changes in the magnetization of ferromagnets in external magnetic fields and mechanical stresses is associated with the need to take into account its total free energy, which depend in a complex way on internal factors (internal and applied stresses, magnetic anisotropy, dislocations, non-equilibrium point defects and inclusions, phase composition). One of the popular mathematical models of magnetic hysteresis of ferromagnetic materials was developed by Jiles and Atherton [1] and then was supplemented by the other scientists. The Jiles-Atherton model (JA-model) has a higher computational performance compared with Preisach phenomenological model and the Stoner-Wolfart hysteresis model [2]. And taking into account modifications, important advantages of the JA-model are the connection with the actual physical parameters of the ferromagnetic material, the calculation of major and any minor hysteresis loops on the same model parameters, the simulation of materials with isotropic and anisotropic properties, the calculation of the effect of mechanical stresses on magnetization.

The hysteresis model for the triaxial (bulk) stress state [3] was based on modifications of the JA-model [4, 5]. However, it turned out to be insufficiently suitable for describing processes in weak magnetic fields.

In the present paper, we refine the model for better describing the magnetization in weak fields, nevertheless maintaining adequate calculations in strong fields.

2. Mathematical formulation

The ideal (but impractical) experiment include triaxial stretching a cube with measuring magnetic characteristics simultaneously. In this paper, the bulk stress state is achieved by combining such types of loading as uniaxial tension/compression, torsion, and internal pressure. Listed types of loadings separately maybe not typical for some structures. They must be considered as the possibility of creating a stress state with independent changes in all three main stresses. The mechanical formulation was described in detail in [6].

Specimens are affected by axial load F , torque T and internal pressure p .

Normal stresses are calculated as:

$$(1) \quad \sigma_z = \frac{F}{\pi(R_{out}^2 - R_{in}^2)},$$

where R_{in} – internal radius of specimen, R_{out} – outer radius of specimen.

The volume-averaged values of the tangential stresses are equal to:

$$(2) \quad \tau = \frac{(R_{out} + R_{in})}{\pi(R_{out}^4 - R_{in}^4)} T.$$

Under the action of internal pressure, tensile circumferential stresses σ_r and compressive radial stresses σ_θ are occurring. Their values, as functions of radius r , can be written as:

$$(3) \quad \begin{cases} \sigma_r = \frac{R_{in}^2}{R_{out}^2 - R_{in}^2} \left(1 - \frac{R_{out}^2}{r^2} \right) p, \\ \sigma_\theta = \frac{R_{in}^2}{R_{out}^2 - R_{in}^2} \left(1 + \frac{R_{out}^2}{r^2} \right) p \end{cases}$$

In the general case, the stress tensor in a cylindrical coordinate system is defined as:

$$(4) \quad A = \begin{pmatrix} \sigma_r & 0 & 0 \\ 0 & \sigma_\theta & \tau \\ 0 & \tau & \sigma_z \end{pmatrix}.$$

By solving the characteristic equation, we find the values of the main stresses averaged over the volume:

$$(5) \quad \begin{cases} \sigma_{j_1} = \frac{1}{R_{out} - R_{in}} \int_{r=R_{in}}^{R_{out}} \sigma_r dr \\ \sigma_{j_2} = \frac{1}{R_{out} - R_{in}} \int_{r=R_{in}}^{R_{out}} \frac{\sigma_\theta + \sigma_z + \sqrt{(\sigma_\theta + \sigma_z)^2 + 4(\tau^2 - \sigma_\theta \sigma_z)}}{2} dr \\ \sigma_{j_3} = \frac{1}{R_{out} - R_{in}} \int_{r=R_{in}}^{R_{out}} \frac{\sigma_\theta + \sigma_z - \sqrt{(\sigma_\theta + \sigma_z)^2 + 4(\tau^2 - \sigma_\theta \sigma_z)}}{2} dr \end{cases},$$

where indexes j_1 , j_2 and j_3 take values 1, 2, 3 from condition $\sigma_1 \geq \sigma_2 \geq \sigma_3$.

To determine the directions of the main stresses, found main stresses are alternately substituted into the system of equations:

$$(6) \quad \begin{cases} (A - E\sigma_j) \vec{n}_j = 0, \\ n_{rj}^2 + n_{\theta j}^2 + n_{zj}^2 = 1 \end{cases},$$

where $\vec{n}_j$ is a vector with components n_{rj} , $n_{\theta j}$ and n_{zj} , ($k = 1..3$), E is identity matrix. In this system of equations, three of the four equations are linearly independent. Thus, having solved three systems, we obtain direction cosines of the main stresses. The angles corresponding to them are denoted by ψ_j .

The JA-model is constructed in two steps: on the first, the anhysteretic magnetization is calculated, then the hysteresis is simulated using a system of differential equations, taking into account changes in the external magnetic field. Anhysteretic magnetization is a magnetization obtained by the simultaneous action of a constant field and an alternating field with an amplitude decreasing to zero. Often, a modified Langevin function is used to describe the anhysteretic magnetization of isotropic materials:

$$(7) \quad M_{an} = M_s \left[\coth\left(\frac{H_e}{a}\right) - \frac{a}{H_e} \right].$$

But it has a singularity at $H_e = 0$ and therefore is not well suited for weak magnetic fields. To find a better function, we carried out an experiment to measure the anhysteretic magnetization curve of structural chromium-nickel steel 15KhN4D.

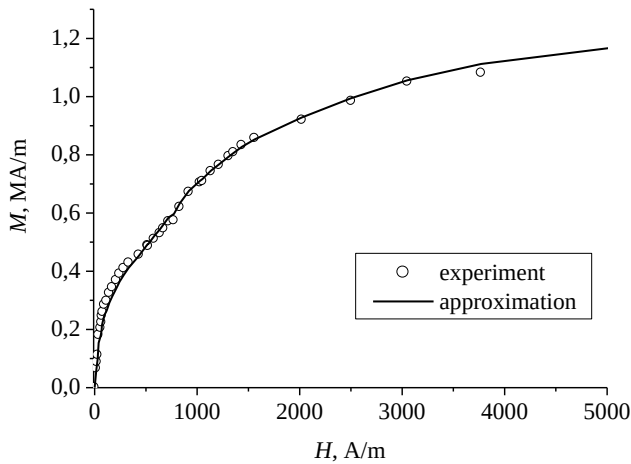


Fig. 1 Anhysteretic magnetization curve and its approximation.

This made it possible to choose a more suitable function that describes anhysteretic magnetization in weak magnetic fields with sufficient accuracy, and at the same time leads it to M_s when $H \rightarrow \infty$:

$$(8) \quad M_{an} = \frac{2}{\pi} M_s \arctan\left(\frac{H_e}{a}\right)$$

M_{an} is a function of effective field H_e :

$$(9) \quad H_e = H \cos \beta + \alpha M + H_\sigma,$$

where β is an angle between the direction of the external field and the vector of the resulting magnetization.

The basic equation of the JA-model, describing the change in total magnetization M , without separating it into reversible and irreversible components can be written as:

$$(10) \quad \frac{dM}{dH} = \frac{c}{1+c} \frac{dM_{an}}{dH} + \frac{1}{1+c} \frac{M_{an}(H) - M(H)}{\delta k - \alpha(M_{an}(H) - M(H))}.$$

Magnetoelastic energy under the action of three orthogonal mechanical stresses consists of three independent parts. The formula proposed previously [3] significantly increases the complexity of calculations because it includes a derivative of magnetostriction which depending on the M . So, as the first approximation for weak fields, we use the more simple form:

$$(11) \quad H_\sigma = \sum_{j=1}^3 \left[(b_{j1} + b_{j2} \sigma_j) \frac{M_{an}(H)}{M_s} (\cos^2 \varphi_j - \nu \sin^2 \varphi_j) \right].$$

Angles φ_j are solutions of three equations:

$$(12) \quad \frac{d}{d\varphi_j} \left[\frac{1}{2} \alpha \mu_0 M^2 - \mu_0 H M \cos(\arccos n_{zj} - \varphi_j) \right] = 0$$

for $j=1..3$.

The angle β between the direction of the external field and the magnetization vector, which is necessary to calculate the effective field (9), can be expressed in terms of the angles that were found above.

$$(13) \quad n_r^* = \sum_{j=1}^3 \cos\left(\frac{\varphi_j}{\psi_j} \arccos(n_{rj})\right)$$

$$(14) \quad n_\theta^* = \sum_{j=1}^3 \cos\left(\frac{\varphi_j}{\psi_j} \arccos(n_{\theta j})\right)$$

$$(15) \quad n_z^* = \sum_{j=1}^3 \cos(\psi_j - \varphi_j)$$

Then angle β is determined by the expression:

$$(16) \quad \beta = \frac{1}{|n^*|} \arccos n_z^*,$$

where $|n^*| = \sqrt{(n_r^*)^2 + (n_\theta^*)^2 + (n_z^*)^2}$. And also we take into account that $\cos \psi_j = n_{zj}$ are known values calculated by (6).

3. Results and discussion

Thus, the simulation of magnetic hysteresis in the bulk stress state must be performed in the following steps:

1. Calculating main stresses (5) and their directions (6).
2. Choosing range of H and the increment dH .

3. Solving the differential equation (10) at each step of H . Considering that anhysteretic magnetization is defined by equation (8), where effective field is given by expressions (9) and (11); angles between the magnetization vector and the directions of action of the main stresses (φ_1 , φ_2 and φ_3) are the solutions of the system of three equations (12); the angle β between the magnetization vector and the direction of external magnetic field is defined by the formula (16).

Experimental studies were conducted on the hollow cylindrical and continuous flat specimens of structural steel 15KhN4D. Permeameter magnetic measurements were made under applied loading. The tests were performed in the elastic strain region. A magnetic field was applied along the axis of the sample.

Although all the parameters of the JA-model have a physical meaning, their experimental determination is not a trivial problem. More practical are the methods of selecting the values of parameters which give a hysteresis loop, with sufficient accuracy reproducing the loop obtained in the experiment. M_s was directly obtained from the experimental major magnetic hysteresis loop. The difference between absolute saturation and the state of technical saturation (in a field of 60 kA/m) can be considered insignificant. To find the optimal parameters, the residuals R are minimized over all unknowns:

$$(17) \quad R = \sum_{i=1}^n (M_i^{\text{exp}} - M(H_i))^2,$$

where (H_i, M_i^{exp}) are experimental points "field-magnetization", n is a number of points in all hysteresis loops. For each set of parameters and for each point, the values $M(H_i)$ are calculated according to the mathematical model described in the previous section.

Analysis of the obtained loops with different stress variations showed that use of parameters b in (11) is not enough. Improving the fit between the experimental and calculated hysteresis loops was received by modifying the parameter k , which was previously considered as a constant:

$$(18) \quad k_\sigma = \left(\xi + (1 - \xi) \exp\left(-\frac{\sigma}{\varsigma}\right) \right) k,$$

where ξ and ς are additional constants for material.

As examples, Figure 2 shows some series of calculated and experimental minor hysteresis loops of the investigated steel. The difference between experimental and calculated values does not

exceed 15%. Figure 2a represents hysteresis loops (only the first quadrant) at various values of the maximum magnetic field. Magnetic fields that are close to the geomagnetic field almost do not create irreversible part of magnetization.

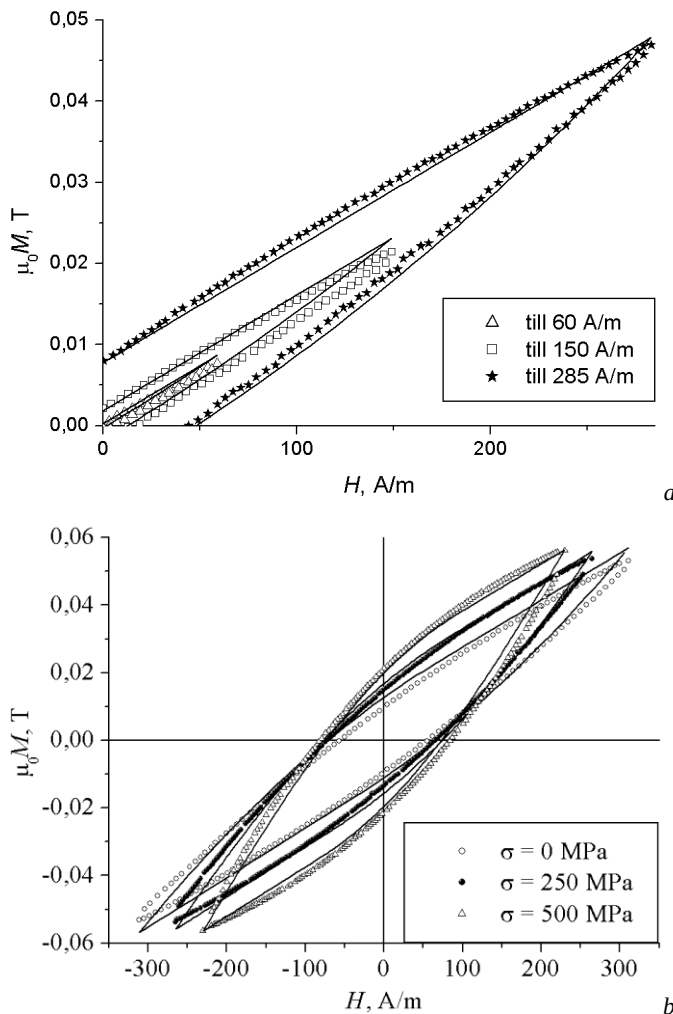


Fig. 2 Comparison of mathematical model and experimentally obtained minor loops. Points from experiments; lines are the model.

Figure 2b shows the effect of mechanical stresses on minor hysteresis loops obtained under magnetization up to the maximum induction of 0.05 T. Tensile stresses lead to the formation of a magnetic texture facilitating magnetization along the specimen axis. The minor loops measured in weak fields grow wider with increasing tensile stresses, their initial portion being steeper.

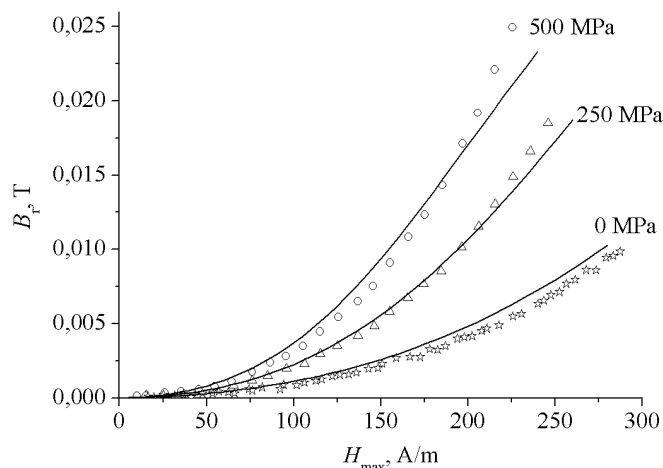


Fig. 3 Residual induction as a function of mechanical stresses and the maximum value of the external magnetic field.

A slight increase in coercive force is also observed. So in a magnetic field of a given strength, growing tensile stress is accompanied by an increase in the relative magnetized volume of the ferromagnetic material.

In details, figure 3 represents changes of residual induction B_r for various minor loops under mechanical stresses. The experiment was carried out on samples of the same grade of steel but from another melting. For this reason, the difference in the integral magnetic characteristics, which are calculated from major hysteresis loops, reaches 10%. In spite of this, for weak magnetic fields, the model rather accurately describes changes of the residual magnetization at the same parameters.

4. Conclusion

A modification of the mathematical model of hysteresis of Jiles-Atherton was proposed for the case of a bulk stress state and weak magnetic fields. It also includes some simplification for increase calculation speed. Comparison of calculated results with experimental data of structural steel showed the adequacy of the model.

5. Acknowledgements

The research was done within the state assignment, theme No. 0391-2016-0005. The experiments were carried out with the use of the equipment of the Plastometriya collective use center affiliated to IES UB RAS.

6. References

- Jiles D. C., Atherton D. L. Theory of ferromagnetic hysteresis // Journal of Applied Physics, 1984, Vol. 55, p. 2115-2120.
- Liorzou F., Phelps B., Atherton D. L. Macroscopic Models of Magnetization // IEEE Transactions on Magnetics, 2000, Vol. 36, № 2, p. 418-428.
- Mushnikov A.N., Putilova E.A. Mathematical Model of Magnetic Hysteresis Under Triaxial Stress State // MATHMODEL'17 Proceedings, p. 55-58.
- Sablik M.J., Rubin S.W., Riley L.A., Jiles D.C., Kaminski D.A., Biner S.B. A model for hysteretic magnetic properties under the application of noncoaxial stress and field. - Journal of Applied Physics, 1993, v. 74, №1, p. 480-488.
- Sablik M.J., Riley L.A., Burkhardt G.L., Kwun H., Cannell P.Y., Watts K.T., Langman R.A. Micromagnetic model for biaxial stress effects on magnetic properties. - Journal of Magnetism and Magnetic Materials, 1994, v. 132, №1-3, p. 131-148.
- Gorkunov E.S., Zadvorkin S.M., Mushnikov A.N., Smirnov S.V., Yakushenko E.I. Effect of Mechanical Stresses on the Magnetic Characteristics of Pipe Steel // Journal of Applied Mechanics and Technical Physics, 2014, Vol. 3, p. 530-538.

MATHEMATICAL MODELLING OF PIEZOELECTRIC DISK TRANSFORMER WITH RING ELECTRODES IN PRIMARY ELECTRICAL CIRCUIT

Dr. Bazilo C., PhD.

Faculty of Electronic Technologies and Robotics – Cherkasy State Technological University, Ukraine
b_constantine@ukr.net

Abstract: Thanks to its unique properties piezoceramics has applications in various fields of engineering and technology. Disk piezoelectric devices are widely used in the elements of information systems. Mathematical model of piezoceramic disk transformer with ring electrode in primary electrical circuit is constructed. An estimate is obtained and calculated and experimental curves of the frequency dependence of the modulus of transformation coefficient of piezoceramic disk transformer with ring electrode in primary electrical circuit are compared in the work. As a result of research of real device's mathematical model a set of geometrical, physical and mechanical and electrical parameters of a real object can be determined which provides realization of technical parameters of piezoelectric functional element specified in technical specifications. The cost of the saved resources is the commercial price of the mathematical model.

Keywords: PIEZOELECTRIC DISK ELEMENT, PHYSICAL PROCESSES, AMPLITUDE-FREQUENCY RESPONSE

1. Introduction

Thanks to its unique properties piezoceramics has applications in various fields of engineering and technology. The relevance of the use of various functional elements of piezoelectronics in radio electronics, information and power systems is explained by their high reliability and small dimensions, which solves the problem of miniaturization of such systems. Piezoelectric disks with surfaces partially covered electrodes are often used to create various functional piezoelectronic devices. Disk piezoelectric devices are widely used in the elements of information systems. In disk piezoelectric elements with surfaces partially covered by electrodes we can simultaneously excite oscillations of compression-tension and transverse bending vibrations. Manipulating the geometric parameters of electrodes and their location relative to each other, you can have a significant effect on the energy of oscillatory motion particular type of material particles of piezoelectric disk volume. It should be especially noted that this piezoelectric element has compatibility with microsystem technology, so it can be made as microelectromechanical structures (MEMS) [1]. One of the main elements of functional piezoelectronics is piezoelectric transformer (PT). Research has shown that PTs can compete with traditional electromagnetic transformers on both efficiency and power density [2-4]. PTs are therefore an interesting field of research [5]. The favorable attributes of the PT are low weight and size and potentially low cost. One additional important characteristic is the high voltage isolation of the ceramic materials used to build PTs [6]. In addition, a piezoelectric transformer is more suitable for mass production than traditional, coil-based transformers [7].

The **purpose** of this paper is to set out the principles of mathematical models construction that are sufficiently adequate to real devices and occurring physical processes using the simplest example of axially symmetric radial oscillations of the piezoelectric disk.

2. Calculation of transformation ratio of piezoelectric transformer with ring electrode in primary electrical circuit

Let us consider a disk piezoelectric transformer (Fig. 1), primary electrical circuit of which consists of electric potential difference generator $U_1 e^{i\omega t}$ (where U_1 is an amplitude value of electric potential difference; $i = \sqrt{-1}$ is an imaginary unit; ω is an angular frequency; t is a time) with output electrical impedance Z_g and ring electrode (position 1 in Fig. 1). The secondary electrical circuit consists of an electrode in the form of a circle (position 2) with connected electronic circuit to it with input electrical impedance Z_n , on which an electric potential difference $U_2 e^{i\omega t}$ is formed. The primary and secondary circuits of piezoelectric transformer do not have a galvanic connection. The energy exchange between the primary and secondary circuits is

carried out by means of axisymmetric radial vibrations of the piezoceramics material particles in the volume of thickness polarized disk (position 3 in Fig. 1).

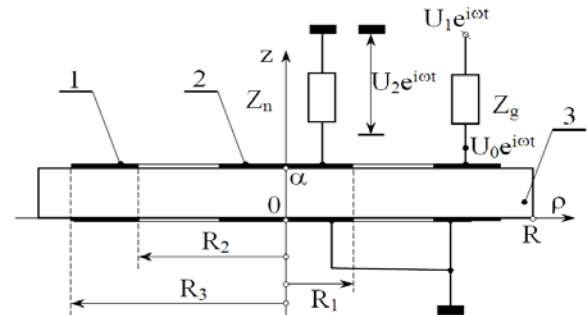


Fig. 1. Calculation scheme of disk piezoelectric transformer

It is obvious that the work of function piezoelectronic element, which is schematically shown in Fig. 1, is fully described by transformation ratio $K(\omega, \Pi) = U_2/U_1$ (Π is a set of geometrical and physical and mechanical properties of the piezoelectric transformer), which is a mathematical model of the device under consideration. Scheme of construction of piezoelectric transformer's mathematical model is outlined in [8].

Methodology for calculating the transformation ratio of piezoelectric transformer with ring electrode in primary electrical circuit is shown in [9]:

$$K(\omega, \Pi) = \frac{U_2}{U_1} = \frac{K_2(\Omega, \Pi)}{1 - i\omega C_0^\sigma Z_g K_3(\Omega, \Pi)}, \quad (1)$$

where

$$K_2(\Omega, \Pi) = \frac{2 f_1(\omega) K_{31}^2 A_2 [J_1(\Omega R_1/R)/(\Omega R_1/R)]}{1 - 2 f_1(\omega) K_{31}^2 A_1 [J_1(\Omega R_1/R)/(\Omega R_1/R)]};$$

$K_{31}^2 = (e_{31}^*)^2 / (c_{11} \chi_{33}^\sigma)$ is a squared electromechanical coupling coefficient for the mode of radial oscillations of thickness polarized piezoceramic disk material particles; e_{31}^* and χ_{33}^σ are material constants and dielectric permittivity for planar stress-strain state of the polarized across the thickness piezoceramic element; c_{11} is a modulus of elasticity for the mode of axially symmetric radial oscillations of the piezoceramic disk material particles; $f_1(\omega)$ is a switching on function or load characteristic of the output ring electrode of the piezoelectric transformer; $f_1(\omega) = \frac{i\Omega^*}{1 - i\Omega^*}$,

$\Omega^* = \omega \tau_n$ is a dimensionless quantity; $\tau_n = C_1^\sigma Z_n$ is a time constant of secondary electrical circuit; $C_1^\sigma = \pi R_1^2 \chi_{33}^\sigma / \alpha$ is a static electrical capacitance in secondary electrical circuit;

$$K_3(\Omega, \Pi) = \frac{2K_{31}^2}{1 - \beta^2} \left\{ [K_2(\Omega, \Pi)A_{41} + A_{42}]J(\Omega) + \right. \\ \left. + [K_2(\Omega, \Pi)A_{51} + A_{52}]N(\Omega) \right\} - 1;$$

$$J(\Omega) = [J_1(\Omega R_3/R) - \beta J_1(\beta \Omega R_3/R)] / (\Omega R_3/R);$$

$$N(\Omega) = [N_1(\Omega R_3/R) - \beta N_1(\beta \Omega R_3/R)] / (\Omega R_3/R);$$

$\beta = R_2/R_3$ is a geometrical parameter of the ring; $J_1(\Omega)$ and $N_1(\Omega)$ are Bessel and Neumann functions of the first order; Ω is a dimensionless wave number; coefficients A_{ij} are defined in [9].

All calculations are performed for piezoceramic disk with radius $R = 33 \cdot 10^{-3}$ m and thickness $\alpha = 3 \cdot 10^{-3}$ m, made of thickness polarized PZT type piezoceramics with following parameters: $\rho_0 = 7400$ kg/m³; $c_{11}^E = 112$ GPa; $c_{12}^E = 62$ GPa; $c_{33}^E = 100$ GPa; $e_{33} = 20$ C/m²; $e_{31} = -9$ C/m²; $\chi_{33}^E = 1800 \chi_0$; $\chi_0 = 8,85 \cdot 10^{-12}$ F/m is a dielectric constant; $Q_M = 100$ is a quality factor of piezoceramics.

In Fig. 2 it is shown the calculated (solid line) and the experimentally obtained (dashed line) curves of the frequency dependence of the modulus of piezoceramic ring-dot disk transformer's transformation coefficient. Naturally, the dimensions of the disk transformer in the calculation and experiment are chosen to be the same, i.e., the radius $R = 33 \cdot 10^{-3}$ m, the thickness $\alpha = 3 \cdot 10^{-3}$ m and $R_1/R = 12/25$, $R_2/R = 15/25$, $R_3/R = 0,999$. The values of the modulus of transformation coefficient of the piezoceramic disk transformer are plotted along the ordinate axis, and the frequency f (dimensionless value $\Omega = \omega \tau_0$, where $\tau_0 = R/\sqrt{c_{11}/\rho_0}$ is a time constant of piezoceramic disk) on the abscissa axis. The frequency $f = 15206$ Hz corresponds to the value $\Omega = 1$.

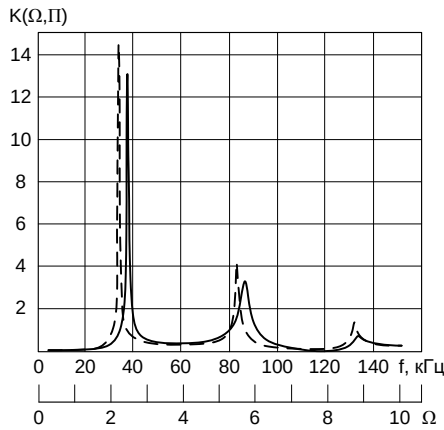


Fig. 2. Calculated (solid line) and experimentally obtained (dashed line) curves of the frequency dependence of the modulus of piezoceramic ring-dot disk transformer's transformation coefficient

When building the model, it was assumed that the thickness of the electrodes located on the surfaces of the disk is very small in comparison with the thickness of the disk α . In other words, the thickness of the electrodes, which, as a rule, does not exceed 15 μ m, was not taken into account for constructing a mathematical model of piezoelectric transformer based on piezoceramic thin disk ($\alpha/R \ll 1$). It should also be noted that mathematical model (1) was built for ring-dot piezoelectric transformer (see Fig. 1) with surfaces partially covered by electrodes (area 1, $\rho \in [0, R_1]$, and area 3, where $\rho \in [R_2, R_3]$) and in the areas where there are no

electrodes (area 2, where $\rho \in [R_1, R_2]$, and area 4, where $\rho \in [R_3, R]$).

As expected, the absolute values of the frequencies of resonances in calculation and experiment differ from each other. So, following the calculation, the frequencies of the first second and third electromechanical resonances are respectively equal to $f_{r1} = 37193$ Hz, $f_{r2} = 88194$ Hz and $f_{r3} = 135330$ Hz; the frequency ratio $\zeta = f_{r2}/f_{r1} = 2.371$.

The experimental values of the same quantities are, respectively, $f_{r1} = 34491$ Hz, $f_{r2} = 83728$ Hz, $f_{r3} = 132325$ Hz and $\zeta = f_{r2}/f_{r1} = 2.428$. If the experimental data are assumed to be true, the error in determining the frequency ratio is $\Delta\zeta = 2.3\%$. The obtained results are explained very simply. The numerical values of the frequencies of resonances s are determined by the dimensions and physicomachanical parameters of the material of disk element. The ratio of the resonances frequencies of the same disk is determined practically only by its dimensions. For this reason, a very satisfactory match between the theoretically and experimentally determined resonance frequency ratios is observed. The discrepancy between the absolute values of the resonance frequencies is explained by the discrepancy between the physicomachanical parameters of the piezoceramics, which were incorporated into the calculation and which are inherent in the experimentally investigated object. Comparing the curves, we can conclude that the quality factor of the material of the experimentally investigated sample is at least 1.2 times larger than included in the quality factor calculation.

3. Discussions and Conclusions

It can be asserted that the character of the variation of both curves, shown in Fig. 2, in a fairly wide frequency range coincides with accuracy to details. This means that the qualitative content of the expression (1) is adequate to the processes that occur in real object. In other words, expression (1) is a mathematical model of piezoelectric ring-dot transformer with ring electrode in primary electrical circuit and sufficiently adequate to the real object and the processes occurring in it. The latter allows us to assume that the mathematical description of the stress-strain state of the disk transformer also corresponds quite well to the real state of things.

Main results of this work can be formulated as follows: mathematical model of piezoelectric trans-former with ring electrode in the primary electrical circuit is constructed; calculated and experimentally obtained curves of frequency dependence of modulus of piezoceramic ring-dot disk transformer's transformation coefficient are estimated and compared.

4. References

- [1] Varadan V., Vinoy K., Jose K. RF MEMS and their applications. Moscow: Technosphaera, 2004, 528 p..
- [2] Bove T., Wolny W., Ringgaard E., and Breboel K. New type of piezoelectric transformer with very high power density. Applications of Ferroelectrics, 2000, Vol. 1, pp. 321-324.
- [3] Flynn A. M. and Sanders S. R. Fundamental limits on energy transfer and circuit considerations for piezoelectric transformers. IEEE Transactions on Power Electronics, 2002, Vol. 17, pp. 8-14.
- [4] Shao W., Chen L., Pan C., Liu Y., and Feng Z. Power density of piezoelectric transformers improved using a contact heat transfer structure. IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, 2012, Vol. 59, pp. 73-81.
- [5] Andersen T., Andersen M. A. E., Thomsen O. C. Simulation of Piezoelectric Transformers with COMSOL. Proceedings of the 2012 COMSOL Conference in Milan. URL:

- https://www.comsol.dk/paper/download/151765/andersen_paper.pdf
- [6] Ivensky G., Shvartsas M., and Ben-Yaakov S. Analysis and Modeling of a Piezoelectric Transformer in High Output Voltage Applications. URL: http://cbucc.com/glecture_file/2002060132032.pdf
 - [7] Huang Y.-T., Wu W.-J., Wang Y.-C., Lee C.-K. Multilayer Modal Actuator-Based Piezoelectric Transformers. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 2007, Vol. 54, No. 2.
 - [8] Tikhonov A.N. Mathematical model. In *Mathematical Encyclopedia*, 1982, Vol. 3, Soviet encyclopedia, Moscow, pp. 574–575.
 - [9] Petrishchev O.N., Bazilo C.V. Mathematical model of piezoelectric disk transformer with ring electrode in primary electrical circuit. Herald of Cherkasy State Technological University, 2016, No. , pp.14–27.

USE OF ICT RESOURCES IN THE HUMANITARIAN SUBJECTS IN BULGARIAN SCHOOLS

M.Sc. Boneva Y.¹, Assist. Prof. PhD. Paunova-Hubenova E.¹, Assist. Prof. Terzieva V.¹, PhD. Dimitrov S.¹
 Institute of Information and Communication Technologies – Bulgarian Academy of Science, Bulgaria¹
 boneva.yordanka@gmail.com, eli_np@hsi.iccs.bas.bg, valia@isdip.bas.bg, sdimitrov85@abv.bg

Abstract: Modern technologies are a part of contemporary life, and they have been part of the educational process in Bulgarian schools for a while. There is no doubt that computer competencies are a necessity, and therefore they are taught as a separate subject at school. However, the use of Information and Communication Technologies (ICT) resources is not limited only to this dedicated subject, these tools also support teaching all other matters. They are used during each stage of school education as well. For exploring the implementations of ICT in an educational environment, an online-based survey aimed at students and teachers was conducted. This paper presents a comparison of the frequency of use of ICT in Bulgarian schools according to different educational stages and subjects with the focus on the humanitarian ones. The current research shows the statistical analysis of the data and a comparison between the answers of students and teachers.

Keywords: SCHOOL EDUCATION, SURVEY-BASED STUDY, ICT RESOURCES, HUMANITARIAN SUBJECTS

1. Introduction

With the introduction of ICT in everyday life, the ordinary teaching methods are increasingly difficult to attract students' attention. The traditional teaching process in the classrooms becomes a real challenge for teachers. It is necessary to integrate new and upgrade the existing technology resources so that they be a tool for the transfer of knowledge, not an obstacle.

There are many studies on the use of information technologies in the Humanitarian classes in different countries. For language learning often used resources are educational videos and games. The authors of [1] present an approach for Computer Assisted Language Learning by means of designing serious games. Another study [2] presents methods for automatic changes of playback speed which can affect the score for learning in the second language. The potentials of using digital educational games for English foreign language learning are discussed in [3]. Many aspects of the usage of technology in foreign language teaching, as well as different types of case studies, are presented in [4].

The authors of [5] discuss the value of hypervideos in presenting historical artifacts in museums. In the framework of the APOGEE project [6], an innovative platform intended for teachers to create adaptive educational maze games in the domain of History is designed. An online educational game in Geography is presented in [7].

According to the number of researches, students can learn from multisensory information quickly and thoroughly. So, for example, something new can be learned much easier with technology-supported lessons than in the traditional way. That is why the authors have started a project that aims to explore the real use of ICTs in teaching practice in Bulgarian schools.

The focus of this research is the implementation of ICT resources in teaching humanitarian subjects in Bulgarian school education. The next section presents the used methodology and the respondents' profile. The results are described in the third section, and their statistic distribution – in the fourth one. The last section presents the analysis of the results and conclusions.

2. Methodology

In this paper, the authors explore the quantitate parameters of the use of ICT resources and tools in the Humanitarian educational context through a quantitative research methodology. The study is based on a comprehensive anonymous online survey that covers many issues, which offer an opportunity for detailed processing and classifying the findings [8]. The survey comprises four different questionnaires – three for students of different age groups and one for teachers [9]. Thanks to the personal efforts of researchers and the cooperation of regional educational authorities the survey was spread via e-mails to the majority of schools in Bulgaria, so the

results can be considered as a national representative. Moreover, the study population covers more than 1600 teachers and 8000 students from all stages of school education in Bulgaria.

In particular, the focus here is on the frequency of usage of different types of innovative pedagogical instruments in each of school stages – primary, low secondary and high secondary. Thus the results from the four different surveys are examined and compared. For in-depth interpretation and analysis of the survey data, the authors use quantitative statistical methods and calculations of indicators such as frequencies, means, and standard deviations. Generally, the developed detailed questionnaires give an opportunity to differentiate the variety of manners for the integration of ICT into the teaching practice. The inclusion of field for comments in each question was of particular importance to gain respondents' opinions and valuable extra information when analyzing data.

For a more accurate analysis of the data, the teachers' profile should be taken into account. In many Bulgarian schools, especially the small ones, the teachers have classes in more than one stage and/ or subject. In such cases, their answers are considered in more than one group, and thus, the sum of the different types of teachers can be higher than the total number of respondents.

The total number of respondent teachers is 1652 from the three school stages, distributed as follows: 605 – primary, 622 – low and 776 – high secondary school teachers. From those the Humanitarian teachers that participate in the survey are 414. Table 1 presents the number of Humanitarian teachers classified according to the school stage and learning subject. The primary teachers are not divided by this criteria, because usually, one person teaches most of the subjects for the earliest stage.

Table 1: Teachers' distribution according to school stage and learning subject.

Teachers in subject \ Stage	Bulgarian language & Literature	Foreign language & Literature	Geography & History	Philosophy
High secondary	74	109	57	12
Low secondary	72	71	71	–
Primary	605 (all subjects)			

3. Results and discussion

The researchers collected and analyzed the data for the use of different types of e-resources by teachers according to the school stage and subject. Here, besides for primary school teachers, the results are also presented for teachers in the humanitarian subjects such as Bulgarian language and literature (BLL), foreign languages

and literature (FLL), History and Geography, and Philosophy. The survey explores what types of ICT resources employ teachers in their practices. The respondents have to choose between five predefined options – very often (almost every day), often (at least ones a week), rarely (at least ones a month), never and do not give answers. The obtained results for different groups of respondents are presented on figures 1 to 7. The missing parts to 100% in the graphics are for those who did not answer (in some cases about half).

Teachers in primary schools

Primary school teachers (see Fig. 1) usually use often and very often presentations (69%), e-textbooks (54%), educational videos (58%) and additional e-resources (37%). All these resources attract pupils' attention easily, so it is worth the efforts to integrate them into lessons. As a whole, there are not many available specialized software and process simulations intended for primary school students, so they are seldom used.

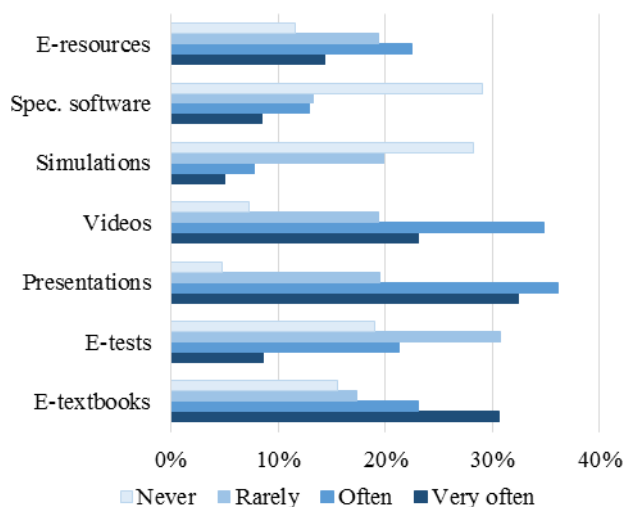


Fig. 1 Primary school teachers – frequency of usage of ICT resources

Teachers in low and high secondary schools

Fig. 2 shows the findings concerning teachers in Bulgarian language and literature in secondary schools. Again, for more than half of respondents, presentations are the most used technology resources, followed by videos and e-textbooks. It should be mentioned that their use prevails in low secondary schools compared to high secondary. E-tests follow the same tendency, but only for about a quarter of teachers.

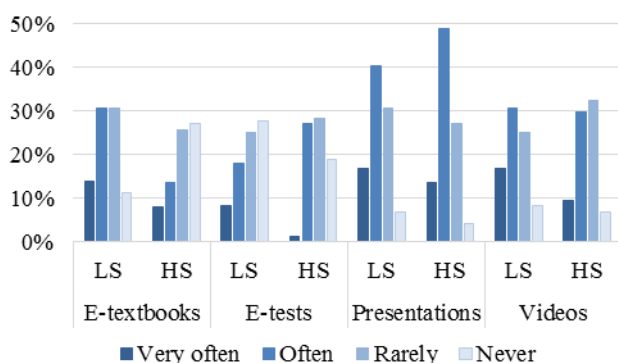


Fig. 2 Low (LS) and High secondary (HS) school teachers in BLL

Almost the same is the situation when considering answers of teachers in foreign languages and literature (Fig. 3). The main difference here is in an almost twice number of teachers who use

very often e-textbooks – 20% and 23% for low and high secondary schools respectively.

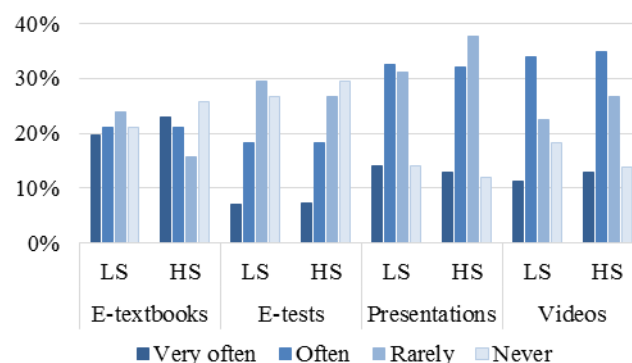


Fig. 3 Low (LS) and High Secondary (HS) school teachers in FLL

The graphics on Fig. 4 compare the frequency of use of simulations, specialized software and other types of e-resources. The latter, usually different kind of web-based resources, are most often employed in language teaching. They are usually used every day by 11% of teachers in a foreign language in high secondary schools and teachers in the Bulgarian language in low secondary schools. Between 16% - 27% of teachers enrich lessons with such resources at least once a week. Simulations and specialized software usually are rarely used (mainly because of their lack), while virtual laboratories are not appropriate in general.

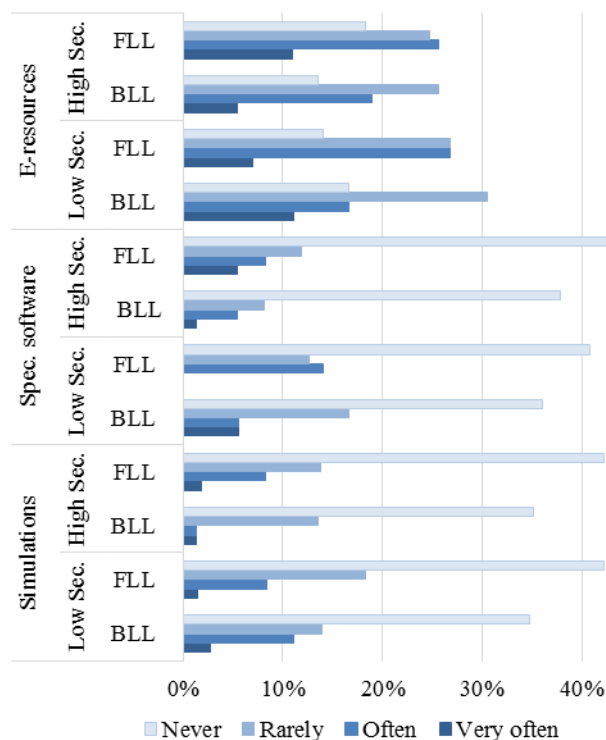


Fig. 4 Low and High Secondary school teachers in BLL and FLL subjects

Generally, in the survey, teachers in foreign languages pointed as usual sources of e-resources the digital versions of textbooks and workbooks that are provided by the publishers as well as various popular electronic dictionaries.

Following the general tendency, as it is shown in Fig. 5, the majority of teachers in Geography and History usually use often and very often presentations, educational videos and e-textbooks. The first two types of resources have more frequent usage in high secondary schools, while the e-textbooks seems to be more popular

in low secondary schools. Only a few of the teachers (11% - 12%) check the students' knowledge by e-tests very often.

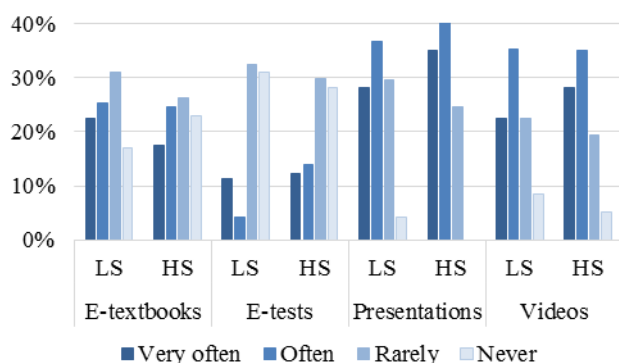


Fig. 5 Low and High Secondary school teachers in Geography and History subjects

The graphics on Fig. 6 shows the comparison by usage frequency of more elaborated technology tools – process simulations, specialized software and different types of e-resources. The latter is quite often integrated into lessons by a third of teachers. Besides no much availability of appropriate learning resources of the other two types, these require additional preparation and efforts to be utilized. Thus they are only used by few teachers in both low and high secondary schools.

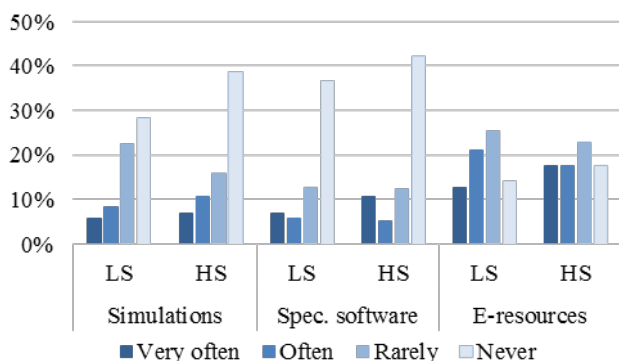


Fig. 6 Low and High Secondary school teachers in Geography and History subjects)

The majority of teachers in Philosophy subjects, like their colleagues, most frequently integrate presentations and educational videos in the lessons. As Fig. 7 shows, about 42% of them use these resources almost every day or every week. Digital textbooks and e-tests present in the pedagogical practice of a third and a quarter of teachers, respectively. The other types of e-resources are not very usual for these learning subjects.

To sum up, in the survey, teachers usually pointed as traditional sources of e-resources the popular platforms with e-lessons such as Ucha.se, Khan Academy, Kahoot, etc.

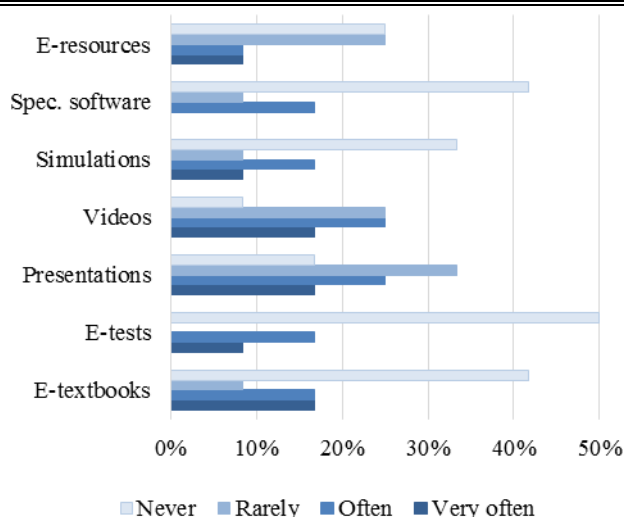


Fig. 7 Low and High Secondary school teachers in Philosophy subjects.

4. Statistical Analysis of Results

Based on the survey findings for usage frequency of different ICT resources, the research explores their appropriateness in teaching humanitarian subjects. The common statistics such as mean values, standard deviation, standard error, difference between M and p-value for the three groups of school subjects are presented in tables 2, 3 and 4. For that purpose the answers are transformed in a 4-level Likert scale from 1 – very rarely/ never use to 4 – very often use (almost every day). These tables compare values corresponding to the answers of low secondary and high secondary teachers.

For Bulgarian language and literature (Table 2) the differences in the means M are statistically significant with a p-value of less than 0.05 only for e-textbooks. The reason is trivial – there are digital versions of textbooks only for students in low secondary classes. The answers for the other ICT-based educational resources are very similar for both low and high secondary school teachers, so the p-value is greater than 0.05.

Table 3 shows that for foreign language teachers the differences are negligible and not statistically significant where the p-value is much greater than 0.05. The same goes for the teachers in geography and history (Table 4). Therefore, according to the teachers, the diverse types of e-resources are equally appropriate for both low and high secondary school students.

The educational games are not very popular resources in all humanitarian subjects in both low and high secondary schools as it is seen from the statistics. Their average use is rare – about once a month, where the most frequently they are used in foreign language teaching and in low secondary Bulgarian language classes.

Table 2: Average frequency of usage of ICT resources in Bulgarian language and literature

Value	Low secondary teachers (N=72)			High secondary teachers (N=74)			Difference	
Answers	M	SD	SE	M	SD	SE	ΔM	P
E-textbooks	2.333333	0.236519	0.118260	1.770270	0.082960	0.041480	0.563063	0.004861
E-tests	1.861111	0.078874	0.039437	1.864865	0.273878	0.136939	-0.003754	0.749578
Presentations	2.611111	0.383771	0.191885	2.648649	0.493012	0.246506	-0.037538	0.649976
Videos	2.361111	0.233544	0.116772	2.202703	0.238194	0.119097	0.158408	0.340418
Simulations	1.444444	0.223952	0.111976	1.202703	0.323249	0.161625	0.241742	0.075947
Virtual Labs	1.125000	0.358212	0.179106	1.054054	0.396467	0.198233	0.070946	0.248554
Spec. software	1.444444	0.216951	0.108476	1.229730	0.317118	0.158559	0.214715	0.175170
E-resources	1.972222	0.074471	0.037235	1.797297	0.136939	0.068469	0.174925	0.827064
Games	1.819444	0.118412	0.059206	1.540541	0.214839	0.107420	0.278904	0.245018

Table 3: Average frequency of usage of ICT resources in foreign languages and literature

Value	Low secondary teachers (N=71)			High secondary teachers (N=109)			Difference	
Answers	M	SD	SE	M	SD	SE	ΔM	P
E-textbooks	2.253521	0.163950	0.081975	2.266055	0.233934	0.116967	-0.012534	0.894641
E-tests	1.873239	0.119251	0.059626	1.853211	0.101645	0.050822	0.020028	0.882591
Presentations	2.380282	0.264694	0.132347	2.403670	0.293184	0.146592	-0.023388	0.878495
Videos	2.239437	0.267304	0.133652	2.348624	0.286321	0.143160	-0.109187	0.572808
Simulations	1.394366	0.240547	0.120274	1.357798	0.255649	0.127825	0.036568	0.901278
Virtual Labs	1.211268	0.330086	0.165043	1.110092	0.367910	0.183955	0.101176	0.199650
Spec. software	1.408451	0.266119	0.133059	1.449541	0.220040	0.110020	-0.041091	0.764238
E-resources	2.014085	0.194747	0.097374	2.091743	0.148216	0.074108	-0.077659	0.954470
Games	1.760563	0.189847	0.094924	1.844037	0.147985	0.073992	-0.083473	0.426237

Table 4: Average frequency of usage of ICT resources in Geography and History subjects

Value	Low secondary teachers (N=71)			High secondary teachers (N=109)			Difference	
Answers	M	SD	SE	M	SD	SE	ΔM	P
E-textbooks	2.492958	0.257861	0.128930	2.280702	0.167128	0.083564	0.212256	0.428567
E-tests	1.746479	0.192345	0.096173	1.947368	0.068370	0.034185	-0.200890	0.380963
Presentations	2.873239	0.437527	0.218764	3.105263	0.562550	0.281275	-0.232024	0.161917
Videos	2.605634	0.343901	0.171950	2.736842	0.411066	0.205533	-0.131208	0.324485
Simulations	1.563380	0.164968	0.082484	1.578947	0.157651	0.078825	-0.015567	0.763189
Virtual Labs	1.225352	0.312151	0.156076	1.228070	0.322182	0.161091	-0.002718	0.812073
Spec. software	1.450704	0.225435	0.112717	1.543860	0.214509	0.107255	-0.093155	0.840805
E-resources	2.056338	0.079985	0.039992	2.105263	0.108148	0.054074	-0.048925	0.916665
Games	1.661972	0.221243	0.110622	1.789474	0.089886	0.044943	-0.127502	0.496901

5. Conclusion

The survey findings show that nowadays when ICTs enter classrooms, they penetrate into the teaching-learning process in many forms not always even have been imagined. The technology can offer quality educational resources, but it is not only because of the technology itself. Teachers need focused knowledge on technology-enhanced pedagogy to make the best of these innovative tools in the classrooms. The educational authorities need to take into account research like the presented here that demonstrates the degree of incorporating a variety of technologies into learning environments so that to create an effective policies to further development in this direction.

The results show that Bulgarian teachers in the humanitarian subjects use mainly educational videos and presentations, followed by e-textbooks and additional e-resources. The least used resources are the virtual labs, specialized software and process simulations which are not appropriate or the teachers do not know about such resources for their subjects.

Acknowledgements

This work has been partly supported by project M02/1 of the Bulgarian National Science fund: "Learning data analytics for ICT resource integration in Bulgarian schools", contract DM 02/1, 13.12.2016.

References

[1] Meyer B. and B. Sørensen, Designing Serious Games for Computer Assisted Language Learning – a Framework for Development and Analysis. Design and Use of Serious Games. Springer, Vol. 37, 2009, pp. 69 - 82

[2] Kishi Y., J. Kim, T. Tachino, Design of Automatic Playback Speed Control System for Learning in Second Language Using Online Videos, in: Proceedings of 10th International Conference EDULEARN18, Palma de Mallorca, Spain, 2nd-4th July 2018, pp. 7023-7029.

[3] Wang X., J. Dostál, Using Digital Educational Games For English Foreign Language Learning, in: Proceedings of 10th International Conference EDULEARN18, Palma de Mallorca, Spain, 2nd-4th July 2018, pp. 144-148.

[4] Motteram, G. (Ed.) Innovations in Learning Technologies for English Language Teaching, London: British Council, 2013.

[5] Rußwurm L., Knowledge Transfer In Museums: Hypervideos As a Way to Present and Communicate Historical Artifacts, in: Proceedings of 10th International Conference EDULEARN18, Palma de Mallorca, Spain, 2nd-4th July 2018, pp. 2767-2776.

[6] Bontchev, B. and Vassileva, D., "Affect-based adaptation of an applied video game for educational purposes", Interactive Technology and Smart Education, Vol. 14, No. 1, 2017, pp. 31-49.

[7] Paunova E., K. Stoilova. Comparative Characteristics of Serious Games. Mathematics and Education in Mathematics, in: Proceedings of the 43rd Spring Conference of the Union of Bulgarian Mathematicians, April 2-6 2014, Bulgaria, pp. 186-191

[8] Terzieva, V., E. Paunova-Hubenova, S. Dimitrov, and Y. Boneva, ICT in STEM Education in Bulgaria, in: Proceedings of 21th International Conference ICL'18, 25-28 September 2018, Kos Island, Greece, pp. 1232 - 1243.

[9] Survey of Teachers in Google Forms (in Bulgarian): <https://docs.google.com/forms/d/1VCOj3r2OPbu7goQTM7QMXf7f20iP5c6Qiv6WF3lBBY>

COMPARISON OF APPROACHES TO ESTIMATION OF TRANSITION MATRIX FOR THE TERRORIST THREAT MARKOV MODEL

Ing. Martin Tejkal, Doc. RNDr. Jaroslav Michálek, CSc.
University of Defence, Brno, Czech Republic
martin.tejkal@unob.cz, jaroslav.michalek@unob.cz

Abstract: Markovian models are often used in modelling a time development of random phenomena. When modelling real world scenarios it is reasonable to assume that the respective phenomena may not be time homogeneous. Based on the sociological and security research, it can be assumed that there is a link between a destabilisation of a society of a given geographical region and the acts of terrorism. This link is utilised in construction of a model for description of the intensity of a terrorist threat based on given determinants/indicators of societal stability. The model is based on the theory of discrete non-homogeneous Markov chains. The theory of generalised linear models (GLMs) is used in the estimation of the probabilities of the categorised level of the terrorist threat. In the contribution the use of different estimates of the categorised level of terrorist threat probabilities is studied. The estimates are determined by GLMs with different input parameters. The influence of the resulting estimate on the transition matrix of the non-homogeneous Markov chain is assessed. Additionally, a real world example utilising the data from Global Terrorism Database of University of Maryland and Organisation for Economical Cooperation and Development is presented.

Keywords: MARKOV CHAIN, NON-HOMOGENEOUS, POISSON GLM, MULTINOMIAL LOGIT GLM, RISK, TERRORISM

1 Introduction

In recent years the mathematical models are finding more and more applications in the field of social sciences. Due to current geopolitical events the ability to asses and predict the risk of terrorism has become topical.

The model discussed in this paper aims to describe the evolution of the categorised level of terrorist threat in time. From mathematical viewpoint this might be seen as a time development of random events. In other fields, mathematical models based on the theory of Markov chains are often used to model random events time development, when some categorisation is involved. The markovian framework provides a good compromise between computational feasibility of the model and taking into account the real world dependencies. The finite or countable state space of the chain is particularly suitable to model the considered categories. Such models are often used in economics and finance. Quite common is the application by banking institutions to model the credit risk, i. e. the risk of default (see [1], [2], [3]).

The problematic of modelling of time development of categorised level of terrorist threat bears some resemblance to the credit risk models. For that reason the framework of Markov chains is utilised in this paper as well.

For some particular events a connection to a set of explanatory variables might be found. For example, it is reasonable to believe that the ability of the companies to fulfil their obligations to the credit institution is influenced by the economical development. Thus, it is reasonable to assume a dependence of the level of the terrorist threat on some economical variables, like gross domestic product etc. Some authors have considered models utilising this additional information provided by the covariates (see [4], [5], [6]).

When it comes to modelling the threat of terrorism, a sociological research suggests, that the level of terrorist threat within a given society (country, etc.) may be connected with an overall stability of the studied society (see [7]). There are numerous indicators of stability of a given society. The authors of this paper have already proposed a model utilising this link connecting the terrorist threat with selected indicators (see [8]). The link was used in the estimation of the transition matrix of the Markov chain. In this paper the research is taken a step forward by providing a different and possibly more suitable way of estimation of the transition matrix as well as a comparison of the two approaches.

2 Prerequisites and means for solving the problem

2.1 Markov Chains

The model presented in this paper is based as well as the model presented in [8] on the theory of Markov chains. The formal description does not differ from the one presented in [8]. For the

reader's convenience, the authors will recall the key points in this subsection.

A sequence $M = \{M_n : n = 0, 1, 2, \dots\}$ of random variables, that attains values from a finite set S is called a Markov chain with finite state space S if

$$(1) \quad \begin{aligned} \mathbb{P}[M_{n+1} = l | M_n = k, M_{n-1} = k_{n-1}, \dots, M_1 = k_1, \\ M_0 = k_0] \\ = \mathbb{P}[M_{n+1} = l | M_n = k]. \end{aligned}$$

for each $l, k, k_{n-1}, \dots, k_0 \in S$. The equation (1) is the well known Markov property.

In the model presented in this paper, the elements $k \in S$ of state space will correspond to the L categories of the level of the terrorist threat.

Out of convenience, henceforth it will not be differentiated between a state of the Markov chain and a category of a level of the terrorist threat. Thus, when we say that the process moved into category k in time step from n to $n + 1$, we actually mean that the process jumped into k -th state in time step from n to $n + 1$ and so on. Furthermore, from the interpretational point of view, the states (categories) denoted by larger values of integers are considered to be the ones of higher level of terrorist threat (i. e. state (category) $k = 3$ denotes higher level of threat than state (category) $k = 1$).

The transition probabilities are defined as follows

$$(2) \quad \pi_{kl}(n, n + 1) = \mathbb{P}[M_{n+1} = l | M_n = k],$$

for each $k, l \in S$. The transition probability is the probability the process jumps into state l (the level of the categorised threat changes to l) in the time step from n to $n + 1$, provided it was in a state k in time n .

Transition probabilities of a homogeneous Markov chain do not depend on time step n , i. e. $\pi_{kl}(n, n + 1) = \pi_{kl}$. In this model, however, it is assumed, that the numbers of terrorist attacks are dependent in each time instant n on a set of m indicator variables $\mathbf{X} = (X_1, \dots, X_m)$. These variables vary in time step n , and hence the notation $\mathbf{X} = \mathbf{X}_n = (X_{n1}, \dots, X_{nm})$ is used. Thus, the resulting Markov chain is non-homogeneous. The transition probabilities of the Markov chain with L states are collected in the transition matrix $\mathbf{P}_{n,n+1}(\mathbf{X}_n)$ of a type $L \times L$ given by

$$(3) \mathbf{P}_{n,n+1}(\mathbf{X}_n) = \begin{pmatrix} \pi_{11}(n, n + 1; \mathbf{X}_n) & \dots & \pi_{1L}(n, n + 1; \mathbf{X}_n) \\ \dots & \dots & \dots \\ \pi_{L1}(n, n + 1; \mathbf{X}_n) & \dots & \pi_{LL}(n, n + 1; \mathbf{X}_n) \end{pmatrix}.$$

The notation $\mathbf{P}_{n,n+1}(\mathbf{X}_n)$ is used to further stress out the fact, that the transition matrix depends on the indicator variables.

The properties of the transition matrix are discussed more in detail in [8] or any classical book on Markov theory (e. g. [9]).

Given the state vector $\mathbf{p}_0 = (p_{1,0}, \dots, p_{L,0})$ in time 0, where $p_{k,0} = \mathbb{P}[M_0 = k]$ for $k \in S$, it is possible to obtain a state vector $\mathbf{p}_n = (p_{1,n}, \dots, p_{L,n})$, where $p_{k,n} = \mathbb{P}[M_n = k]$ for some $n \in \mathbb{N}_0$ in the following way

$$(4) \quad \mathbf{p}_n = \mathbf{p}_0 \mathbf{P}_{0,1}(\mathbf{X}_1) \mathbf{P}_{1,2}(\mathbf{X}_2) \dots \mathbf{P}_{n-2,n-1}(\mathbf{X}_{n-1}) \mathbf{P}_{n-1,n}(\mathbf{X}_n).$$

The transition matrices $\mathbf{P}_{0,1}(\mathbf{X}_1), \dots, \mathbf{P}_{n-1,n}(\mathbf{X}_n)$ need to be estimated.

2.2 Multinomial Response Generalised Linear Models

Generalised linear model (GLM) is used to connect the categorised level of terrorist threat with the indicator variables. The GLM is used to obtain an estimate of the transition matrix row by row. In the paper [8], a Poisson GLM is utilised. Here the use of the multinomial response generalised linear models is proposed (see [10], [11]).

Let M_n for $n \in \mathbb{N}$ be a nominal scale response variable with L categories. The multicategorical model used in this paper to obtain the estimates of the transition probabilities in k -th category is the multinomial logit model (see [10], [11]), that is given by

$$(5) \quad \log \left(\frac{\mathbb{P}[M_n=l]}{\mathbb{P}[M_n=L]} \right) = \beta_{l0,k} + \mathbf{x}_{n,\bullet,k} \boldsymbol{\beta}_{l,k},$$

where $\beta_{l0,k}$ and $\boldsymbol{\beta}_{l,k} = (\beta_{l1,k}, \dots, \beta_{l(L-1),k})$ are unknown parameters, which correspond to state $k = 1, \dots, L-1$. For the baseline category $l = L$, set $\beta_{L0,k} = 0$ and $\boldsymbol{\beta}_{L,k} = \mathbf{0}$. The $\mathbf{x}_{n,\bullet}$ are the values of the indicator variables $\mathbf{X}_n$ at the time step n .

3 Solution of the examined problem

The multinomial logit GLM was used repeatedly in sequence for each state $k \in \{1, \dots, L\}$ the process visited. In this way a vector $\hat{\pi}_{k,\bullet} = (\hat{\pi}_{k1}, \dots, \hat{\pi}_{kL})$ of probability estimates was obtained, such that the respective probability estimates correspond to the k -th row of the desired transition matrix estimate.

3.1 Transition Matrix Estimation Using Multinomial Logit GLM

The multinomial logit GLM described in the Subsection 2.2 is computed for each category $k \in \{1, \dots, L\}$. Hence, for each category $k \in \{1, \dots, L\}$, the estimate $\hat{\boldsymbol{\beta}}_k = (\hat{\beta}_{10,k}, \hat{\boldsymbol{\beta}}_{1,k}^T, \dots, \hat{\beta}_{L0,k}, \hat{\boldsymbol{\beta}}_{L,k}^T)$ of the vector of parameters of the GLM is obtained.

For the multinomial logit GLM for each category $k \in \{1, \dots, L\}$ the estimates of $L-1$ fitted values $\hat{\pi}_{k1}(\mathbf{x}_{n,\bullet}), \dots, \hat{\pi}_{k(L-1)}(\mathbf{x}_{n,\bullet})$, given the vector of covariates $\mathbf{x}_{n,\bullet} = (x_{n1}, \dots, x_{nm})$ for some $n \in \mathbb{N}$ are obtained through a formula

$$(6) \quad \hat{\pi}_{kl}(\mathbf{x}_{n,\bullet}) = \frac{\exp(\hat{\beta}_{l0,k} + \mathbf{x}_{n,\bullet,k} \hat{\boldsymbol{\beta}}_{l,k})}{1 + \sum_{r=1}^{L-1} \exp(\hat{\beta}_{r0,k} + \mathbf{x}_{n,\bullet,k} \hat{\boldsymbol{\beta}}_{r,k})},$$

where $l = 1, \dots, L-1$. For $l = L$ we then have

$$(7) \quad \hat{\pi}_{kL}(\mathbf{x}_{n,\bullet}) = 1 - \sum_{r=1}^{L-1} \hat{\pi}_{kr}.$$

The estimate of the probability of transition from state k to state l in time instant from n to $n+1$ for $n \in \{1, \dots, s\}$ is then defined as $\hat{\pi}_{kl}(n, n+1; \mathbf{X}_n) = \hat{\pi}_{kl}(\mathbf{x}_{(n+1),\bullet})$.

3.2 Transition Matrix Estimation Using Poisson GLM

The detailed description of the method of obtaining the estimates of the transition matrix using the Poisson GLM is provided in the paper

[8]. The key points will be recalled here.

The response in this case are the numbers of terrorist attacks. These are split into categories together with the corresponding observations of explanatory variables, based on whether the value of the observation of the response lies between a selected bounds of the respected category.

The Poisson GLM is given by

$$(8) \quad \log(\lambda_n) = \beta_0 + \mathbf{x}_n \cdot \boldsymbol{\beta},$$

where the expectation λ_n (the average number of terrorists attacks in the given category) depends on unknown parameters $\beta_0, \boldsymbol{\beta}$ and on the values $\mathbf{x}_n$ of the indicator indicator variables $\mathbf{X}_n$ at the time step n .

This model is again applied for each category $k = 1, \dots, L$ in order to obtain vector of estimates $\hat{\alpha}_k = (\hat{\beta}_{0,k}, \hat{\boldsymbol{\beta}}_k^T)$ of the parameters $\alpha_k = (\beta_{0,k}, \boldsymbol{\beta}_k^T)$.

The estimate $\hat{\lambda}_{nk}$ of the parameter λ_{nk} for each category $k = 1, \dots, L$ is then obtained via the formula

$$(9) \quad \hat{\lambda}_{nk} = e^{\hat{\beta}_{0,k} + \mathbf{x}_n \cdot \hat{\boldsymbol{\beta}}_k}.$$

Then a random variable Z_k with Poisson distribution with the value $\hat{\lambda}_k$ as an expectation parameter is defined for each category $k = 1, \dots, L$. The transition transition from a category k into a category l , $k, l \in S$ is then defined as the probability of the random variable Z_k attaining values between the lower and upper bound of the category.

4 Results and discussion

4.1 Real Data Example

The approaches proposed in this paper and in paper [8] were applied to estimate a transition matrix using a real dataset. The dataset consisted of quarterly numbers of terrorist attacks in France between the years 1981 and 2018 obtained from the Global Terrorism Database of University of Maryland, and two indicator explanatory variables, the gross domestic product and the unemployment rate obtained from the Organisation for Economical Development. Detailed description of the dataset is provided in [8]. For the readers convenience the authors will present figures of the dataset in question.

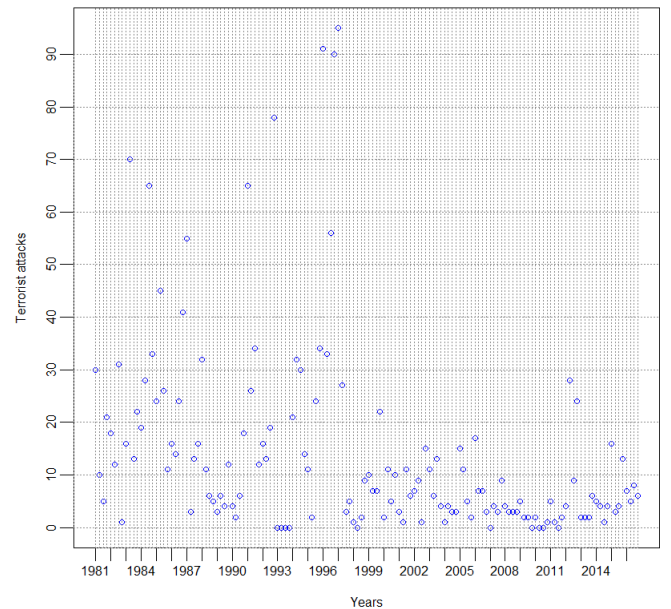


Fig. 1: Time series of quarterly observations of the numbers of terrorist attacks



Fig. 2: Time series of quarterly observations of GDP in constant prices of US Dollar (left horizontal axis) and of unemployment rate in percentage of labour force (right horizontal axis)

From the Fig 2 one can see that, the gross domestic product grew in average over time, while the unemployment rate stagnated or increased slightly. This development might be considered still more or less optimistic scenario. It should be noted, that there might be a relation between the two explanatory variables (see Fig 2). The last available observation of the explanatory variables in the year 2016 was taken as the “future” observation.

Categorisation of the data was carried out in the following manner. Into the first category were placed observations greater or equal to 0 and lesser than 6, into the second observations greater or equal to 6 and lesser than 13. Into the third category were placed observations greater or equal to 13. The transition matrix estimate obtained using the Poisson GLM described in [8] is the following:

$$(10) \quad \hat{\mathbf{P}}_{0,1;Po}(\mathbf{X}) = \begin{pmatrix} 0.990 & 0.010 & 0.000 \\ 0.335 & 0.636 & 0.029 \\ 0.000 & 0.007 & 0.993 \end{pmatrix}.$$

The matrix obtained using the multinomial logit GLM described in this paper is

$$(11) \quad \hat{\mathbf{P}}_{0,1;Mult}(\mathbf{X}) = \begin{pmatrix} 0.812 & 0.095 & 0.093 \\ 0.361 & 0.530 & 0.109 \\ 0.457 & 0.453 & 0.090 \end{pmatrix}.$$

The transition matrix estimated using the multinomial logit GLM reflects the optimistic development of responses, favouring the category with the least terrorist threat. This is not entirely true about the transition matrix obtained via the Poisson GLM described in [8]. In the estimate (10) the probabilities $\hat{\pi}_{kk}(0,1)$ on the diagonal attain large values, while the further we move in the row away from the diagonal entry, the lower the estimated probability (compare $\hat{\pi}_{12}(0,1)$ and $\hat{\pi}_{13}(0,1)$ in (10). Additionally, the probability $\hat{\pi}_{33}(0,1)$ of staying in the third category of the highest terrorist threat seems to be unreasonably high, given the optimistic scenario. In [8] it was

suggested, that this might rather be due to the properties of the Poisson GLM itself, and does not reflect the reality very well.

Note, that the size of real input dataset is rather small, in addition the observations of terrorist attacks contain a large number of very small observations and a small set of relatively high values (see Fig. 1). This, altogether with the particular choice of the bounds of the categories further worsens the effect that the Poisson model has on the estimates of the entries on the diagonal.

The problem of unreasonably high values on the diagonal seems not to be present in the estimate (11). The probability is more evenly distributed among the respective categories, while at the same time reflecting the positive development scenario. For the first and the third row we have $\hat{\pi}_{kl} > \hat{\pi}_{kr}$ for k fixed and $l < r$. The exception is the second row, where the probability jumping from the second state back to the second state is the highest.

The link between the response observations and the covariates was weak when modelled via the Poisson GLM. Only for the category $k = 3$ the tests for each parameter rejected the null hypothesis $H_0: \hat{\beta}_{j,k} = 0$ for $j = 0,1,2$. In the case of multinomial logit GLM the link between the response and the covariates was weak as well. Additionally, no matter the choice of the baseline category, a separation occurred for some logits (see [10]), rendering the tests of significance for the respective parameters impossible.

5 Conclusion

The paper [8] presents a non-homogeneous Markov chain model for prediction of a categorised level of terrorist threat. The non-homogeneity of the Markov chain stems from the assumption that the categorised level of the terrorist threat and thus the transition matrix of the chain is dependent on a set of indicators of destabilisation of a society, that vary in time. A Poisson GLM is used to connect the categorised level of terrorist threat with the indicator covariates and estimate the transition matrix row by row. This paper proposes a modified approach using multinomial logit GLM instead of the Poisson one. A comparison of the two approaches is carried out on a real dataset.

The estimates of the transition probabilities on the diagonal of the transition matrix obtained via the Poisson GLM tend to attain large values. This is assumed to be rather a property of the Poisson GLM, that does not reflect the real situation. Especially unexpected is the large value of the probability of staying in the category of the largest terrorist threat in case of an optimistic development scenario. This effect is not present, when the multinomial logit GLM is applied instead. The transition probabilities tend to be more evenly distributed among the respective categories, while reflecting the character of the scenario (for the rather optimistic scenario given by the real data the transition to the categories of lower threat tend to be higher than the ones to the categories of higher threat).

It should be noted, however, that in general the connection between the responses and the covariates regardless of the used GLMs was weak. Additionally, the size of the currently available real data sample proved to be insufficient. The proposed multinomial logit model did not share the diagonal values-increasing effect of the Poisson model, and thus seems to be more suitable, in further research however, it would be necessary to choose different set of indicators to further study the model. Additionally, a way to generate the datasets for a thorough simulations study would need to be developed. In order to increase the sample size a panel data approach may be considered. That is however, beyond the scope of this paper.

References

- [1] Jarrow, R.A., Lando, D., Turnbull, S.M., A Markov model for the term structure of credit risk spreads. The Review of Financial Studies 10 (2), 1997, 481–523.

[2] Altman, E. I. and Kao, D. The implications of corporate bond rating drift. *Financial Analysts Journal*, 48(3), 1992, 64–75.

[3] Carty, L. V. and Fons, J. S. Measuring Changes in Corporate Credit Quality, *Journal of Fixed Income*, 4(1), 1994, 27–41.

[4] Nickell, P., Perraudin, W. and Varotto, S. Stability of rating transitions. *Journal of Banking & Finance*, 24(1–2), 2000, 203–227.

[5] Koopman, S. J. and Lucas, A. Business and default cycles for credit risk. *Journal of Applied Econometrics*, 20(2), 2005, 311–323.

[6] Gavalas, D. and Syriopoulos, T. Bank Credit Risk Management and Rating Migration Analysis on the Business Cycle. *International Journal of Financial Studies*, 2(1), 2014, 122–143.

[7] Svoboda, I. and Hrbata, M. Extremism and Terrorism as

Destabilizing Factors of Society, *Vojenské rozhledy*, 23 (55), N. 1, p., 2014, 33 - 41 (In Czech)

[8] Tejkal, M. and Michálek, J. Markov Model for Description of Intensity of Terrorist Threat. Abstracts of XXXII International Conference Problems of Decision Making under Uncertainties (PDMU-2018) , July 3-9, 2018, Prague, Czech Republic.

[9] Resnick, S. I., *Adventures in stochastic processes*, Boston: Birkhäuser, 1992.

[10] Agresti, A., *Categorical data analysis*, 3rd ed. Hoboken, NJ: Wiley, 2013.

[11] Fahrmeir, L. and Tutz G. *Multivariate statistical modelling based on generalized linear models*. New York, Springer-Verlag, 1994.

CLASSIFICATION OF PROTEIN STRUCTURES BY USING FUZZY KNN CLASSIFIER AND PROTEIN VOXEL-BASED DESCRIPTOR

Prof. Dr. Mirceva G., Prof. Dr. Naumoski A., Prof. Dr. Kulakov A.

Faculty of computer science and engineering, Ss. Cyril and Methodius University in Skopje, Skopje, R. Macedonia

georginamirceva@gmail.com

Abstract: Protein classification is among the main themes in bioinformatics, for the reason that it helps understand the protein molecules. By classifying the protein structures, the evolutionary relations between them can be discovered. The knowledge for protein structures and the functions that they might have could be used to regulate the processes in organisms, which is made by developing medications for different diseases. In the literature, plethora of methods for protein classification are offered, including manual, automatic or semiautomatic methods. The manual methods are considered as precise, but their main problem is that they are time consuming, hence by using them a large number of protein structures stay uncategorized. Therefore, the researchers intensively work on developing methods that would afford classification of protein structures in automatic way with acceptable precision. In this paper, we propose an approach for classifying protein structures. Our protein voxel-based descriptor is used to describe the features of protein structures. For classification of unclassified protein structures, we use a k nearest neighbors classifier based on fuzzy logic. For evaluation, we use knowledge for the classification of protein structures in the SCOP database. We provide some results from the evaluation of our approach. The results show that the proposed approach provide accurate classification of protein structures with reasonable speed.

Keywords: PROTEIN STRUCTURE, PROTEIN CLASSIFICATION, PROTEIN VOXEL-BASED DESCRIPTOR, K NEAREST NEIGHBORS, FUZZY LOGIC

1. Introduction

Bioinformatics community intensively analyze protein molecules for the reason that they are essential in the organisms. The processes in the organisms can be controlled by the interactions of proteins. The knowledge gathered from the examination of proteins may be used for drug design, where the functions of the proteins in these interactions is taken into consideration. Using various types of techniques, the structures of the protein molecules have been examined. The information about protein structures is stored in the Protein Data Bank (PDB) [1], [2]. Due to the fast improvements in these techniques, the structures of proteins are determined with fast speed. However, the methods that provide classification of proteins are not able to classify them with the same speed that leads to gap in the number of proteins with determined structures and the number of proteins that are classified. Thus, there is a great necessity for developing methods for classification of protein structures.

The current literature offers various methods for classification of proteins. In the SCOP (Structural Classification Of Proteins) method [3] the decision is done in manual way, where the experts visually examine the proteins. However, the manual methods are time consuming and are not able to follow the speed of determining novel protein structures. Therefore, there are also automatic methods and semiautomatic methods. For example, the CATH (Class, Architecture, Topology and Homologous superfamily) method [4] tries to classify proteins in automatic manner first, and if the decision could not be made, then manual decision is performed.

Some methods align the sequences of the protein structures, also known as primary structures, in order to perform classification of the proteins. The most known methods in this group are Needleman–Wunch [5], BLAST [6] and PSI-BLAST [7]. However, the protein sequence is a particular sequence of amino acid residues that folds in some specific way in the three-dimensional space. In that way, two amino acid residues could be close in the space, while far in the protein sequence. Therefore, the methods based on alignment of protein sequences are not able to recognize distant homology between proteins. For that purpose, also there is a group of methods that analyze the tertiary structures of proteins, like CE [8], MAMMOTH [9] and DALI [10]. Third group of methods, like SCOPmap [11] and FastSCOP [12], combines both sequence and structure alignment.

Besides alignment of the sequences or structures of proteins, feature vectors could be extracted, and then the proteins can be compared based on the distance between their feature vectors. In the

literature, there are methods that use features of the sequence [13] or structure [14] of proteins, as well as both of them. By extracting the vector of features, the protein is presented by a point in the feature space. Later, in the classification stage, the amount of information that is processed is significantly lower than by making direct alignment of proteins, so the time needed for classification is much lower. For that purpose, in our earlier study [15], we presented feature vectors that contain features that represent the protein structures. These feature vectors could be used as inputs and by using some classification method, a prediction model could be generated.

In this paper, we use the protein voxel-based descriptor presented in [15] and we apply a fuzzy k nearest neighbors classification method [16] to classify the unknown structures.

Here is an outline of the structure of the remaining of this paper. Section 2 provides description of the proposed approach, where the protein voxel-based descriptor [15] and fuzzy k nearest neighbors classification method [16] are presented. The results of the evaluation of the approach are presented and discussed in Section 3, whereas Section 4 presents the main conclusions and points out directions for further advancements of the approach.

2. The Proposed Approach

The approach used in this study has two steps. The first step performs extraction of the feature vectors of the training protein structures. In this study, we use the protein-voxel based descriptor [15]. After mapping the proteins in the feature space, next, in the second step we use the fuzzy k nearest neighbors classification method [16] to determine the class in which a given query protein would belong to.

Protein Voxel-Based Descriptor

The protein voxel-based descriptor contains features of the primary, secondary and tertiary structure of the protein. Regarding the tertiary structure, we use the voxel descriptor [17] that is originally proposed for comparing 3D objects. Concerning the primary and secondary structure, the features are extracted as in [18].

The extraction of the protein voxel-based descriptor is done in the following way. First, a mesh model of the protein structure is generated, by making triangulation of the atoms of the protein that are treated as spheres and with triangulation they are presented by a given number of triangles. In order to obtain feature vector that is invariant to translation, the protein is translated so that its center of

mass is in the center of the coordinate system. With the aim to obtain feature vector that is invariant to scaling, next we perform scaling of the mesh model thus the most distant vertex of the triangles in the mesh model is at a distance equal to 1 from the center of mass.

After obtaining the mesh model, next, a voxelization is made. With this step, we transform the continual into discrete space. For that purpose, first, discretization is made, where the continuous three-dimensional space is divided into equal cubes named voxels. Then, for each of the voxels, we calculate the ratio of the area of the mesh that is in the inspected voxel. For that purpose, the triangles are divided into p_j^2 triangles with a surface $\delta = S_j / p_j^2$, where S_j denotes the area of the triangle T_j that is currently divided. If a given triangle T_j has vertices in one voxel, then p_j is set to 1. Otherwise, it is calculated as

$$(1) \quad p_j = \left\lceil \sqrt{p_{\min} \frac{S_j}{S}} \right\rceil,$$

where S is the total surface of the triangles and $p_{\min}$ defines the quality of the approximation. In this study, we use $p_{\min} = 32000$ as in [17]. For each voxel, the surface that is placed in the voxel is incremented for δ .

With the previous step, as output, we obtain three-dimensional matrix that could be used as feature vector. However, the number of features contained in this matrix could be significantly reduced. Moreover, this three-dimensional matrix as a feature vector is not invariant to rotation. Therefore, in the next step, we apply 3D Discrete Fourier Transform thus obtaining feature vector that is invariant to rotation. In this way, the new version of the feature vector is also a three-dimensional matrix.

Next, the indices are shifted so the voxel in the center has indices (0, 0, 0). Because there is a symmetry between the elements of the obtained feature vector, therefore only the non-symmetrical features are considered that corresponds to the values of the voxels with indices that satisfy $1 \leq |p| + |q| + |s| \leq N/2$, where (p, q, s) are the indices of the voxel and N is the number of slices for one coordinate used in the discretization of the space. Further, the features are divided by the feature that correspond to the voxel with indices (0, 0, 0). More details about the extraction of the geometrical features contained in the voxel descriptor can be found in [17] and [15].

Besides the features of the tertiary structure of the proteins, we also consider several features of their primary and secondary structure. In this study, we use the features used in [18]. Regarding primary structure, we consider the ratio of each amino acid and the ratio of the hydrophobic amino acids in the protein. From the features of the secondary structure, we consider the ratios of the types of helices, as well as the number of occurrences of each type of secondary structure element (helix, sheet and turn). More details about these features can be found in [18] and [15].

Fuzzy k Nearest Neighbors Classifier

For classifying a given query sample (protein chain in this case), first its protein voxel-based descriptor is extracted. Then, this query sample is compared with the training samples and its class is determined by using the fuzzy k nearest neighbors (Fuzzy KNN) classification method [16].

The Fuzzy KNN method is inspired from the well known k nearest neighbors (KNN) classification method [19], but adjusted for sets in fuzzy logic. KNN could be used to perform simple majority voting of the nearest neighbors, or the neighbors may have different weights in the voting in order to give higher significance to the votes of the closer neighbors. For the second approach, the distance between the examined sample and the nearest neighbor could be used in order to calculate the weight of the vote of that neighbor.

Let assume we are using k nearest neighbors for making decisions. The Fuzzy KNN method first identifies the k nearest neighbors for the inspected sample q , which are denoted as NN (nearest neighbors). For that purpose, the similarity between a given training sample x and the query sample q is calculated as $S(x, q) = 1/D(x, q)^2$, where $D(x, q)$ denotes the distance between x and q . Then, the examined sample q is classified by maximizing

$$(2) \quad \frac{\sum_{x \in NN} S(x, q) M_c(x)}{\sum_{x \in NN} S(x, q)},$$

where c is the examined class, while $M_c(x)$ denotes the membership function for that class. The membership function could be crisp, defined as

$$(3) \quad M_c(x) = \begin{cases} 1, & x \in C \\ 0, & x \notin C \end{cases},$$

where C is a set of the samples in class c . In this study, since we are using an approach based on fuzzy set theory, therefore instead of using a crisp function we are using the gradual function presented in [16]. The membership function that is used is defined as

$$(4) \quad M_c(x) = \begin{cases} 0.51 + 0.49 \frac{n_c}{k}, & x \in C \\ 0.49 \frac{n_c}{k}, & x \notin C \end{cases},$$

where $n_c = |C|$ is the size of the set C .

3. Results and Discussion

For evaluation, we used 6145 protein chains that are classified in 150 different SCOP domains. The protein chains correspond to the samples in the set, while the domains are the output classes (the possible values for the target attribute). The information about the classification of the protein chains in SCOP domains is obtained from the SCOP database [3]. The distribution of the protein chains used in this study is approximately uniform. This set is divided into training set (90% of the chains) and test set (10% of the chains). In this way, the training set that is obtained has 5531 chains, while the remaining 614 chains form the test set. As evaluation measure, the classification accuracy is used, which gives evidence about the percent of the test samples that are classified correctly. The experimental results are presented on Fig. 1. We made experiments by using different values for the number of nearest neighbors that are considered for making predictions ($k = 1, 2, 3, 4, 5, 10, 15, 20$). In order to give better picture of the benefit of using fuzzy sets instead of classical sets, we made experiments by applying the classical KNN classification method and the Fuzzy KNN classification method that is based on fuzzy logic.

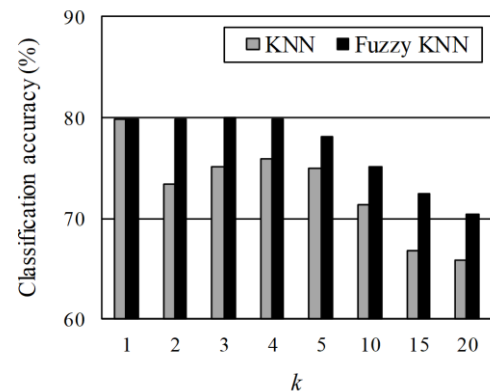


Fig. 1 Classification accuracy achieved by KNN and Fuzzy KNN classification methods by using different number of nearest neighbors k .

As it can be seen from the results, Fuzzy KNN provides better results than the classical KNN classification method. By using Fuzzy KNN, it is best to use between $k=1$ and $k=4$ nearest neighbors, and then by increasing the number of nearest neighbors that are considered for making decisions, the classification accuracy declines. Regarding KNN, the best result is obtained by using $k=1$ nearest neighbor. By using $k=2$ and $k=3$, lower accuracy is obtained than by using $k=4$. Then, by increasing the number of nearest neighbors (for $k>4$), the classification accuracy declines. The highest classification accuracy of 79.97% is achieved with Fuzzy KNN classifier by considering $k=3$ nearest neighbors.

The obtained results for the proposed approach are comparable to the results obtained with the existing approaches. Also, this approach provides classification of proteins with reasonable speed. Namely, the time needed for classification of all 614 test protein chains used in this study is in range of minutes, which is much better than the time needed with manual methods.

4. Conclusion

In this study, we proposed a novel approach that could be used to make decisions about the classification of protein chains into SCOP domains. For each training protein chain, its protein voxel-based descriptor is extracted, which is a feature vector that contains features about the primary, secondary and tertiary structure of the protein. The classification of an unknown test protein chain is done in two steps. First, its protein voxel-based descriptor is extracted, and then by applying the Fuzzy KNN classifier the class of the unknown protein chain is determined.

For evaluation, we used information about the classification of the protein chains in SCOP domains. The results show that it is best to use up to $k=4$ nearest neighbors, while by further growth of the number of nearest neighbors the results are getting worse. The Fuzzy KNN classifier was compared with the classical KNN method, and the results indicate that Fuzzy KNN is better.

As future work, we plan to extend this study in several directions. Regarding the feature vector, besides the protein voxel-based descriptor, we also plan to use some of the other feature vectors that we already used in our previous studies where these descriptors were used for retrieving similar protein structures. Concerning the classification method, we may also apply some other distance and similarity measures for estimating the similarity between two samples. Besides the Fuzzy KNN classification method, we plan to apply other classification methods, including methods based on classical set theory as well as fuzzy set theory.

Acknowledgment

This work was partially financed by the Faculty of Computer Science and Engineering at the "Ss. Cyril and Methodius University in Skopje", Skopje, R. Macedonia.

References

- [1] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, "The protein data bank," *Nucleic Acids Res.*, vol. 28, no. 1, pp. 235–242, 2000.
- [2] RCSB Protein Data Bank, <http://www.rcsb.org>, 2018.
- [3] A. G. Murzin, S. E. Brenner, T. Hubbard, and C. Chothia, "Scop: a structural classification of proteins database for the investigation of sequences and structures," *J. Mol. Biol.*, vol. 247, no. 4, pp. 536–540, 1995.
- [4] C. A. Orengo, A. D. Michie, D. T. Jones, M. B. Swindells, and J. M. Thornton, "CATH – a hierarchic classification of protein domain structures," *Structure*, vol. 5, no. 8, pp. 1093–1108, 1997.
- [5] S. B. Needleman and C. D. Wunsch, "A general method applicable to the search for similarities in the amino acid sequence of two proteins," *J. Mol. Biol.*, vol. 48, no. 3, pp. 443–453, 1970.
- [6] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *J. Mol. Biol.*, vol. 215, no. 3, pp. 403–410, 1990.

- [7] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Res.*, vol. 25, no. 17, pp. 3389–3402, 1997.
- [8] I. N. Shindyalov and P. E. Bourne, "Protein structure alignment by incremental combinatorial extension (CE) of the optimal path," *Protein Eng.*, vol. 11, no. 9, pp. 739–747, 1998.
- [9] A. R. Ortiz, C. E. Strauss, and O. Olmea, "Mammoth: an automated method for model comparison," *Protein Sci.*, vol. 11, no. 11, pp. 2606–2621, 2002.
- [10] L. Holm and C. Sander, "Protein structure comparison by alignment of distance matrices," *J. Mol. Biol.*, vol. 233, no. 1, pp. 123–138, 1993.
- [11] S. Cheek, Y. Qi, S. S. Krishna, L. N. Kinch, and N. V. Grishin, "SCOPmap: automated assignment of protein structures to evolutionary superfamilies," *BMC Bioinformatics*, vol. 5, pp. 197–221, 2004.
- [12] C. H. Tung and J. M. Yang, "fastSCOP: a fast web server for recognizing protein structural domains and SCOP superfamilies," *Nucleic Acids Res.*, vol. 35, W438–W443, 2007.
- [13] K. Marsolo, S. Parthasarathy, and C. Ding, "A Multi-Level Approach to SCOP Fold Recognition," *IEEE Symposium on Bioinformatics and Bioeng.*, pp. 57–64, 2005.
- [14] P. H. Chi, Efficient protein tertiary structure retrievals and classifications using content based comparison algorithms, PhD thesis, University of Missouri-Columbia, 2007.
- [15] G. Mirceva, I. Cingovska, Z. Dimov, and D. Davcev, "Efficient approaches for retrieving protein tertiary structures," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 9, no. 4, pp. 1166–1179, 2012.
- [16] J. M. Keller, M. R. Gray, and J. R. Givens, "A fuzzy k-nearest neighbor algorithm," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 15, no. 4, pp. 580–585, 1985.
- [17] D. V. Vranic, 3D Model Retrieval, Ph.D. Thesis, University of Leipzig, 2004.
- [18] P. Daras, D. Zarpalas, A. Axenopoulos, D. Tzovaras, and M. G. Strintzis, "Three-Dimensional Shape-Structure Comparison Method for Protein Classification," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 3, no. 3, pp. 193–207, 2006.
- [19] D. Aha, D. Kibler, "Instance-based learning algorithms," *Machine Learning*, vol. 6, no. 1, pp. 37–66, 1991.

DETERMINATION OF SURFACTANT EFFECTIVE DIFFUSION COEFFICIENT USING INVERSE PROBLEM FOR WARD-TORDAI EQUATION

M.Sc. Eng. Stasiak M.¹, M.Sc. Eng. Andrzejewski A.², Prof. Prochaska K.²

Institute of Mathematics, Faculty of Electrical Engineering, Poznan University of Technology, Poland¹

Institute of Chemical Technology and Engineering, Faculty of Chemical Technology, Poznan University of Technology, Poland²

marcin.stasiak@put.poznan.pl

Abstract: The main purpose of the study was to assess the usability of a new calculation algorithm for determining the surfactants effective diffusion coefficient. The adsorption of a mixture of non-ionic surfactants onto flat water-air interface was considered. Presented algorithm is composed of a two parts: Ward-Tordai equation solver based on Nyström method for integral equations and golden ratio optimization method used in inverse problem. In the investigation the Langmuir model of an isotherm was assumed. Presented algorithm was successfully used to determine the non-ionic surfactant - Nafol 810D (BRENNTAG) effective diffusion coefficient in the diffusion controlled adsorption process.

Keywords: DIFFUSION COEFFICIENT, DIFFUSION, ADSORPTION, INTEGRAL EQUATION, INVERSE PROBLEM

1. Introduction

In modelling of mass transport processes an important role plays the diffusion process, which describes the movement of molecules from a domain of higher concentration to a region of lower concentration. The mass flux is caused by a gradient of concentration of a particle in a solution.

Adsorption is defined as the adhesion of atoms, ions or molecules from a gas, liquid or dissolved solid to a surface. This process, which is a consequence of surface energy, creates a film of the adsorbate on the surface of the adsorbent. In the systems with surfactants the monolayer is usually created, however the effects of electrostatic interactions are neglected [1]. Adsorption is usually described through the isotherms, that is, the amount of adsorbate on the adsorbent as a function of its pressure (if gas) or concentration (if liquid) at constant temperature. These functions describes the relationship between the bulk concentration $c \left[\frac{\text{mol}}{\text{m}^3} \right]$ and surface excess $\Gamma \left[\frac{\text{mol}}{\text{m}^2} \right]$ of an adsorbed compound. Many models of isotherms were proposed in the literature. Their level of complexity depends on the phenomena according to the considered model. The easiest way to describe the above mentioned relation is the linear correlation called Henry isotherm and is given by eq. 1 [2].

$$\Gamma = K_H c \quad (1)$$

where K_H is a Henry constant. It is a huge simplification of the real problem because it doesn't include the finiteness of the surface. It might be useful for modelling the adsorption process in the terms of the low concentration of surfactant. More compatible with the real processes is the non-linear Langmuir isotherm given by eq. 2 [2].

$$\Gamma = \Gamma_\infty \frac{K_L c}{1 + K_L c} \quad (2)$$

where K_L is Langmuir constant and $\Gamma_\infty \left[\frac{\text{mol}}{\text{m}^2} \right]$ is a maximum surface excess. Both of the presented isotherms are connected with a dynamics of adsorption process. However the Gibbs isotherm, given by eq. 3, is useful to determine the equilibrium value of surface excess $\Gamma_{eq} \left[\frac{\text{mol}}{\text{m}^2} \right]$ of the considered system

$$\Gamma = - \frac{1}{RT} \frac{d\gamma}{d \ln c} \quad (3)$$

where $R \left[\frac{\text{J}}{\text{mol K}} \right]$ is a gas constant, $T \text{ [K]}$ is a temperature of the system and $\gamma \left[\frac{\text{mN}}{\text{m}} \right]$ is a surface tension of the system [2].

Adsorption of non-ionic surfactants is characterized by two mechanisms – diffusion of molecules to the interface and adsorption on the surface (see fig. 1).

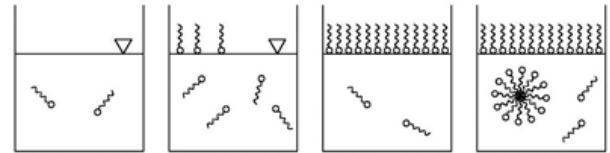


Fig.1. Diffusion and adsorption at the interface

When diffusion process is slower than adsorption process, what is a frequent situation in the real systems, one can say that the adsorption of surfactant is controlled by diffusion. This paper is focused on determination of surfactant effective diffusion coefficient which describes the rate of the whole process. The model of above presented process is given by one-dimensional

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} \quad (4)$$

diffusion partial differential eq. 4 [3,4].

$$D \frac{\partial c}{\partial x} \Big|_{x=0} = \frac{\partial \Gamma}{\partial t} \quad (5)$$

$$\lim_{x \rightarrow \infty} c(x, t) = c_b \quad (6)$$

with boundary conditions (5) and (6):

$$c(x, t)|_{t=0} = c_b \quad (7)$$

for $t > 0$, where c_b is a bulk concentration, and initial condition (7):

The solution of diffusion model which describes dynamic of surfactant adsorption for a long time processes was given by Ward

$$\Gamma(t) = \sqrt{\frac{D}{\pi}} \left\{ 2c_b \sqrt{t} - \int_0^t \frac{c(\Gamma(\tau))}{\sqrt{t-\tau}} d\tau \right\} \quad (8)$$

and Tordai in 1946 by eq. 8 [5].

where $D \left[\frac{\text{m}^2}{\text{s}} \right]$ is a diffusion coefficient, $c(\Gamma(t))$ is the relations between concentration and surface excess and t is a time. Eq. 8 is a Volterra integral equation of the second kind. Linearity or non-linearity of this equation depends on chosen model of isotherm of adsorption. Linear case with Henry isotherm was solved

analytically by Sutherland [6]. In this paper the Langmuir isotherm will be used. The non-linear case doesn't have the analytical solution.

2. Results and discussion

The diffusion of the commercial surfactant was investigated (trade name Nafol 810D). The experimental data was given by the Du Nöuy ring method [7]. The fig. 2 shows the change of surface tension during the experiment of the samples of different bulk concentration.

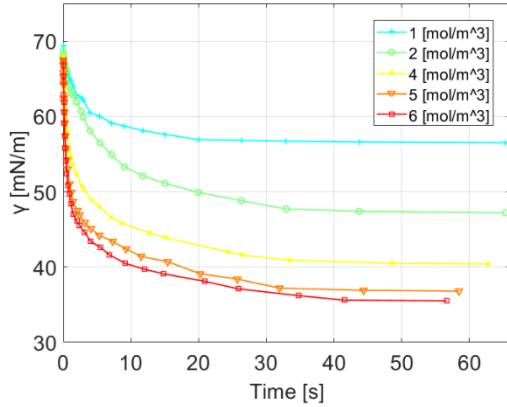


Fig.2. The change of surface tension during the experiment of the samples of different bulk concentration

The Szyszkowski isotherm, given by eq. 9 [8], was fitted to the experimental relation between the surface tension and the bulk concentration of the compound (fig. 3).

$$\gamma = \gamma_0 \left(1 - B_{sz} \ln \left(\frac{c}{A_{sz}} + 1 \right) \right) \quad (9)$$

where γ_0 is the water surface tension, γ is the surface tension associated with the surfactant concentration c , A_{sz} and B_{sz} are Szyszkowski constants, which have the physical meanings. A_{sz} describes the measure of tendency to interfacial adsorption and B_{sz} characterizes the orientation of the adsorbed molecule. Eq. 3 and 9 were used to determine the values of Γ_{eq} and Γ_{∞} of the considered system.

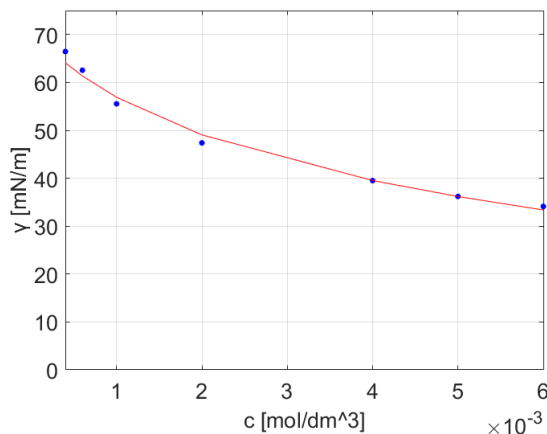


Fig.3. Experimental data of the equilibrium surface tension depending on the bulk concentration

The relation between surface excess and surface tension is given by eq. 10 [4,9]. It was used to determine the surface excess Γ as a function of time.

$$\Gamma(t) = \Gamma_{\infty} \left(1 - \exp \left(\frac{\gamma(t) - \gamma_0}{RT\Gamma_{\infty}} \right) \right) \quad (10)$$

The main purpose of the study was to assess the usability of a new calculation algorithm for determining the surfactants effective diffusion coefficient using the inverse problem for Ward-Tordai equation.

The diffusion coefficient was investigated in the interval $10^{-7} \div 10^{-16} \left[\frac{m^2}{s} \right]$. In order to solve the nonlinear Ward-Tordai integral equation associated with Langmuir isotherm the Nystrom method was used [10,11]. Nystrom method is a discrete method of solving integral equations. In short, the time interval is discretized into finite vector with constant time steps and the integral in eq. 8 is approximated by chosen quadrature. In a consider case the trapezium rule of integration was used. The final discretized form of eq. 8 is given by eq. 11 with initial step $\Gamma(0) = 0$, which means that any of surfactant molecule isn't adsorbed on the surface in the beginning of the diffusion process.

The golden ratio method was used in a optimization part of the inverse problem algorithm [12]. The aim function given by eq. 12 is the maximum norm of the subtraction of the experimental and numerical data vectors, respectively Γ_{num} and Γ_{exp} .

$$\Psi(D) = \|\Gamma_{num} - \Gamma_{exp}\|_{\infty} \quad (11)$$

The experimental data was compared with numerical data and error was minimized successfully in a numerical inverse problem algorithm for all considered samples. Fig. 4 shows the experimental and numerical data of the surface excess in the concentration bulk

$$\Gamma(t_i) = 2c_b \sqrt{\frac{Dt_i}{\pi}} - \frac{1}{K_L} \frac{D}{\pi} \sum_{j=1}^i A_j \frac{\Gamma(t_j)}{\sqrt{t_i - t_j} (\Gamma_{\infty} - \Gamma(t_j))} \quad (12)$$

5 $\frac{mol}{m^3}$ case.

for $i = 2, 3, \dots$

In every time iteration the nonlinear eq. 12 has to be solved. Eq. 12 can be reformulated into quadratic eq. which discriminant is a positive number. Only one solution of this eq. has a physical meaning.

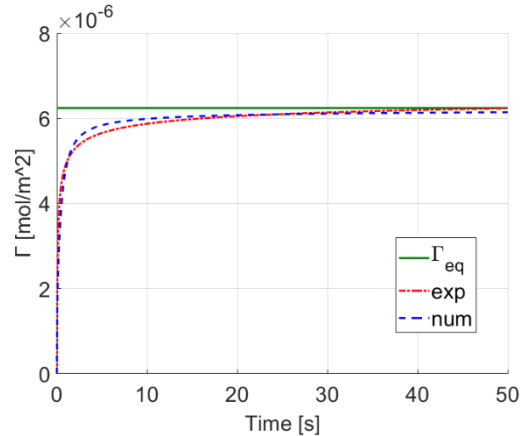


Fig.4. Experimental and numerical data of the surface excess in the bulk concentration 5 $\frac{mol}{m^3}$

Fig. 5 shows the minimization of the aim function during the optimization process. Fig. 6 shows reaching the optimal value of effective diffusion coefficient during the optimization process. Both figures are associated with the 5 $\frac{mol}{m^3}$ case as an example. The results of the first 15 iteration are showed. The optimal value for all investigated cases was reached after 30 ÷ 40 iterations.

The dependence between the effective diffusion coefficient and the concentration of the investigated compound is presented in the table 1.

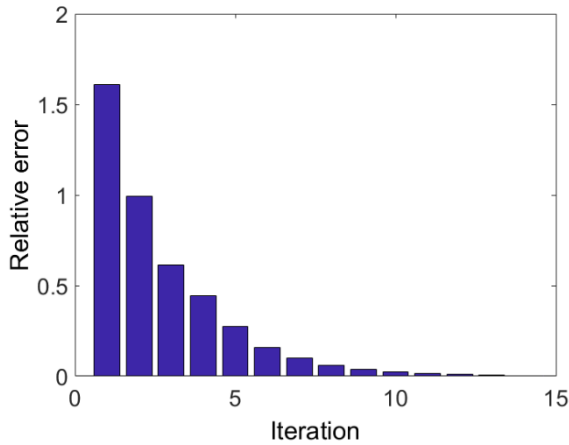


Fig.5. Experimental and numerical data of the surface excess in the bulk concentration $5 \frac{\text{mol}}{\text{m}^3}$

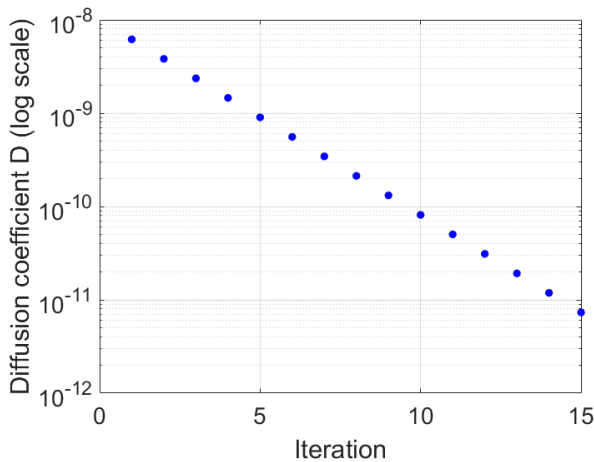


Fig.6. Experimental and numerical data of the surface excess in the bulk concentration $5 \frac{\text{mol}}{\text{m}^3}$

Table 1: Values of effective diffusion D coefficient for different bulk concentration c_b

Bulk concentration $c_b \frac{\text{mol}}{\text{m}^3}$	Effective diffusion coefficient $D \frac{\text{m}^2}{\text{s}}$
1	$2,87 \cdot 10^{-11}$
2	$3,34 \cdot 10^{-12}$
4	$1,62 \cdot 10^{-12}$
5	$1,68 \cdot 10^{-12}$
6	$1,14 \cdot 10^{-12}$

3. Conclusions

Above presented procedure was used to determine the effective diffusion coefficient of commercial surfactant Nafol 810D. In every case of the bulk concentration presented algorithm minimized the aim function and the inverse Ward-Tordai problem was solved. Result of the paper is the effective diffusion coefficient dependence on the bulk concentration of the surfactant.

Presented algorithm can be successfully use to determine the effective diffusion coefficient of amphiphilic surfactants based on the experimental data of dynamic of adsorption process.

4. Acknowledgments

This work was supported by the Faculty of Electrical Engineering, Poznan University of Technology, under grant no. 04/43/DSMK/010.

5. References

- [1] Chattoraj D. K., Birdi K. S., Adsorption and the Gibbs Surface Excess. New York and London, Plenum Press, 1984.
- [2] Rosen M. J., Kunjappu J. T., Surfactants and Interfacial Phenomena: Fourth Edition, New York, John Wiley & Sons Ltd., 2012.
- [3] Crank J., The Mathematics of Diffusion, Oxford, Clarendon Press, 1975.
- [4] Miller R., Joos P., Fainerman V. B., Dynamic surface and interfacial tensions of surfactant and polymer solutions, Adv. Colloid Interface Sci., vol. 49, no. C, 1994, p. 249–302.
- [5] Ward A. F. H. and Tordai L., Time-dependence of boundary tensions of solutions I. The role of diffusion in time-effects, J. Chem. Phys., vol. 14, no. 7, 1946, p. 453–461.
- [6] Sutherland K. L., Australian Journal of Scientific Research, Series A: Physical Sciences, vol. 5, 1952, p.683.
- [7] Gorevski N., Miller R., Ferri J. K., Non-equilibrium exchange kinetics in sequential non-ionic surfactant adsorption: Theory and experiment, Colloids Surfaces A Physicochem. Eng. Asp., vol. 323, no. 1–3, 2008, p. 12–18.
- [8] Radzio K., Prochaska K., Interfacial activity of trioctylamine in hydrocarbon/water systems with nonorganic electrolytes, J. Colloid Interface Sci., vol. 233, no. 2, 2001, p. 211–218.
- [9] Miller R., Aksenenko E. V., and Fainerman V. B., Dynamic interfacial tension of surfactant solutions, Adv. Colloid Interface Sci., vol. 247, 2017, p. 115–129.
- [10] Delves L.M, Mohamed J.L., Computational methods for integral equations, New York, Cambridge University Press, 1992.
- [11] Li X., Shaw R., Evans G. M., Stevenson P., A simple numerical solution to the Ward-Tordai equation for the adsorption of non-ionic surfactants, Comput. Chem. Eng., vol. 34, no. 2, 2010, p. 146–153.
- [12] Chapra S., Applied Numerical Methods with Matlab for Engineers and Scientist, Third Edition, New York, McGraw-Hill, 2008.