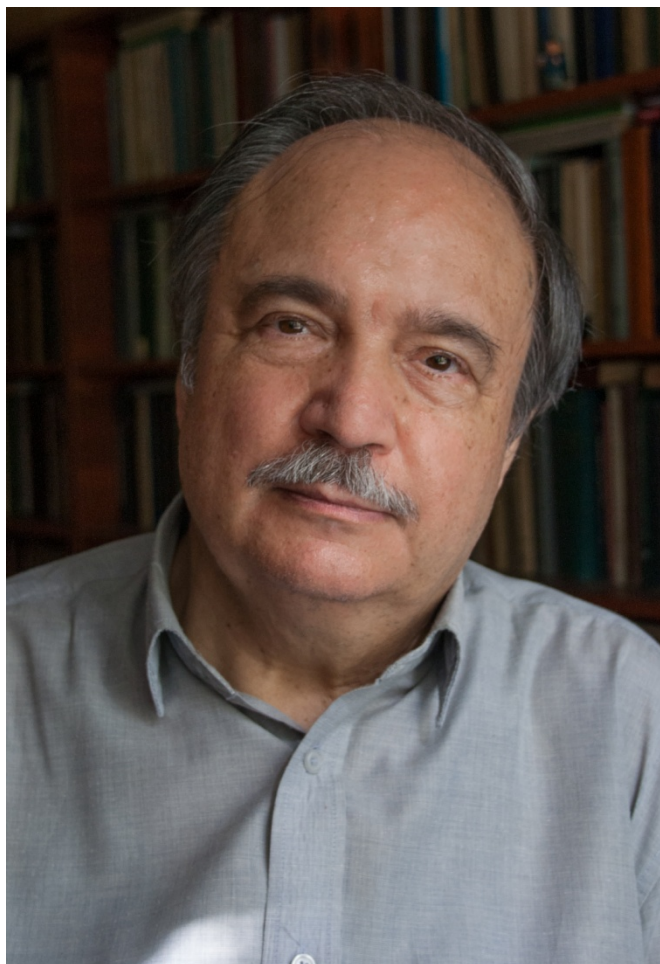


INTERNATIONAL SCIENTIFIC JOURNAL

MATHEMATICAL MODELING

**YEAR I ISSUE 2/2017 ISSN
(PRINT) 2535-0986
ISSN (WEB) 2603-2929**

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



Уважаемые коллеги!

Задача нашего журнала – предоставить возможность широкой научной общественности познакомиться с новыми идеями и исследованиями теоретического и прикладного характера из разных областей научного творчества, но объединенных общим языком описания – языком математики. Использование этого языка позволяет лучше осознать единство и взаимозависимость различных, на первый взгляд, явлений и процессов. Мы надеемся, что наш журнал поможет молодым исследователям научиться грамотно излагать свои мысли, а также сможет подсказать им новые направления для творческого поиска. Мы надеемся также, что журнал сможет стать ареной содержательных профессиональных дискуссий, ибо, как хорошо известно, «в споре рождается истина». Со своей стороны, мы сделаем все от нас

зависящее, чтобы создать для авторов комфортные и объективные условия общения с Редакционной коллегией. Успехов!

Dear colleagues!

Our journal is aimed to provide an opportunity for the wide-range scientific community to get acquainted with new ideas in theoretic and applied research, in various areas of scientific creativity, united by a common language - the language of mathematics. This language is able to clarify an understanding of the unity and interdependence of phenomena and processes, which initially are seemed to be different. We hope that our journal could be helpful for young researchers, who want to learn how to express correctly their thoughts, and also show them new directions of creativity. We hope that our journal will become an arena for meaningful professional discussions, cause, as it's well known, "Discussio mater veritas est". For our part, we will do our utmost to create comfortable and objective conditions for communication of authors with the Editorial Board. Good luck!

Andrei Firsov

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
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. Mendeleev 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 Găvrută, 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 I, ISSUE 2 / 2017

EDITOR IN CHIEF:

Prof. D.Sc. ANDREY FIRSOV

Peter the Great St.Petersburg Polytechnic University, Russia

C O N T E N T S

THEORETICAL FOUNDATIONS AND SPECIFICITY OF MATHEMATICAL MODELLING

EVOLUTIONARY MATHEMATICAL MODELS WITH DISTRIBUTED PARAMETERS ON THE NET AND NETLIKE DOMAIN

Prof. Dr.Math. Zhabko A., Prof. Dr.Math. Provorotov V. 72

FUNDAMENTALS OF THE KINETIC THEORY OF MULTICOMPONENT EMULSIONS

Prof., Dr. Tech. Sci. Firsov A.N. 75

CONCEPTUAL CYBERNETIC MODEL OF TEACHING AND LEARNING

Prof. Ing. Veronika Stoffová, CSc. 80

MATHEMATICAL MODELLING OF TECHNOLOGICAL PROCESSES AND SYSTEMS

MODELING AND SIMULATION OF INDUSTRIAL PROCESSES

Christo Boyadjev 84

ROTATING OF A BALL IN CHAMBER FILLED WITH A FLUID

Prof., Dr. Dementev O. 88

SOFTWARE ASSURANCE OF THE SYNTHESIS AND DESIGN OF HYPERBOLOID GEAR DRIVES

Assoc. Prof. Abadjieva E. PhD., Prof. Sc. D. Abadjiev V. PhD. 92

THE EFFECT OF OPTICAL AND RECOMBINATION LOSSES IN $\text{Cu}_2\text{ZnSn}(\text{S,SE})_4$ -BASED THIN-FILM SOLAR CELLS WITH CDS, ZNSE, ZNS WINDOW AND ITO, ZNO CHARGE-COLLECTING LAYERS

M.Sc. Dobrozhan O.A., St. Danylchenko P.S., Ph.D. Grynenko, V.V., Prof. Dr. Opanasyuk A.S. 96

MATHEMATICAL MODELLING OF SOCIO-ECONOMIC PROCESSES AND SYSTEMS

INTEGRAL ASSESSMENT OF ENVIRONMENTAL QUALITY AND THE QUALITY OF LIFE OF THE POPULATION OF THE ARCTIC REGIONS OF RUSSIA IN THE PERIOD FROM 2003 TO 2015

Prof. Dr. V.V. Dmitriev 100

TRENDS IN DATA ANALYSIS: STATE, DEVELOPMENT PROSPECTS

Doctor in Economic sciences, Prof., I. A. Katsko; PhD in Economic sciences, P. Yu. Velichko Student, M. Nikogda 105

MATHEMATICAL MODELLING OF SOCIO-ECONOMIC PROCESSES AND SYSTEMS

EXACT RECONSTRUCTION VERSION OF RADON TRANSFORMATION IN TOMOSYNTHESIS

Morgun O., Nemchenko K., Vaisburd A., Viktynska T. 110

APPLICATION OF FUZZY MODELING TO PREDICT THE DISEASE OF STAFF FROM EXPOSURE TO WORKING CONDITIONS

Klimova I. V., Smirnov Yu. G. 113

SIMULATION MODELING OF AUDITORY FUNCTION

Ass. Prof., Dr. Eng. Donii O., Dr. Med. Pisanko V., Ass. Prof., Dr. Eng. Kulinich A., Ass. Kotliar S. 117

EVOLUTIONARY MATHEMATICAL MODELS WITH DISTRIBUTED PARAMETERS ON THE NET AND NETLIKE DOMAIN

ЭВОЛЮЦИОННЫЕ МАТЕМАТИЧЕСКИЕ МОДЕЛИ С РАСПРЕДЕЛЕННЫМИ ПАРАМЕТРАМИ НА СЕТЯХ И СЕТЕПОДОБНЫХ ОБЛАСТЯХ

Prof. Dr.Math. Zhabko A. , Saint-Petersburg State University, Department of control theory, Faculty of Applied Mathematics – Control Process, Saint-Petersburg, Russian Federation;

e-mail: zhabko.apmath.spbu@mail.ru

Prof. Dr.Math. Provorotov V., Voronezh State University, Department of partial differential equations and probability theory, Faculty of Mathematics, Voronezh, Russian Federation; e-mail:

wwprov@mail.ru

Abstract: Mathematical models of evolutionary processes on the network and setepodobnoj area. The method, which applies to many tasks of optimal control of differential systems, the status of which is defined by the weak solutions of evolutionary equations of mathematical physics on networks and setepodobnyh areas. This method is very common and is applicable to a wide class of linear tasks that have an interesting analogy with also multi-phase tasks of mechanics (in particular, the theory of plasticity). The results obtained in this manner for a specific equation with distributed parameters in the setepodobnoj area, serve not only to demonstrate the method, but also of interest to applications.

Keywords: PARABOLIC EQUATION, DISTRIBUTED PARAMETERS, NETLIKE DOMAIN, WEAK SOLUTIONS, OPTIMAL CONTROL

1. Introduction

Mathematical models of evolutionary processes on the net and the netlike domain. The method, which applies to many tasks of optimal control of differential systems, the status of which is defined by the weak solutions of evolutionary equations of mathematical physics on the net and the netlike domain. This method is very common and is applicable to a wide class of linear tasks that have an interesting analogy with also multi-phase tasks of mechanics (in particular, the theory of plasticity). The results obtained in this manner for a specific equation with distributed parameters on the net or netlike domain, not only serves to demonstrate the method, but also of interest to applications.

2. Results and discussion

The study consists of two parts – an analysis of the evolution mathematical models with spatial variable that changes to the net (limited graf) and on the netlike domain. While each case is put and the optimal control problem is studied.

2.1. Mathematical model with distributed parameters on the net (graph)

We denote (see [1, 2]): Γ – limited-oriented geometric graph with edges γ which are parametrized by segment $[0,1]$; $\partial\Gamma$ and $J(\Gamma)$ – sets of boundary and internal nodes of the graph, respectively; Γ_0 – the union of all the edges that do not contain the endpoints; $\Gamma_t = \Gamma_0 \times (0,t)$, $\partial\Gamma_t = \partial\Gamma_0 \times (0,t)$ ($t \in [0,T]$).

Necessary space and sets: $C(\Gamma)$ – the space of continuous on Γ functions; $L_p(\Gamma)$ ($p=1,2$) --- Banach space measurable functions on the Γ_0 , summable with a p -th

degree (spaces $L_p(\Gamma_T)$ are defined similarly); $L_{2,1}(\Gamma_T)$ –

the space of continuous from $L_1(\Gamma_T)$ with norm

$$\text{PuP}_{L_{2,1}(\Gamma_T)} = \int_0^T \left(\int_{\Gamma} u^2(x,t) dx \right)^{1/2} dt; \quad W_2^1(\Gamma) - \text{the}$$

space of continuous from $L_2(\Gamma)$, that have a generalized derivative of the 1st order also from the space $L_2(\Gamma)$;

$W_2^{1,0}(\Gamma_T)$ – the space of continuous from $L_2(\Gamma_T)$, that have a generalized derivative of the 1st order of the variable x , that belongs $L_2(\Gamma_T)$ (similarly we introduce the space $W^1(\Gamma_T)$).

Пусть $V_2(\Gamma_T)$ – the set of all functions $u(x,t) \in W_2^{1,0}(\Gamma_T)$

with finite norm $\text{PuP}_{2,\Gamma} = \max_{0 \leq t \leq T} \|u(\cdot, t)\|_{L_2(\Gamma)} + \|u_x\|_{L_2(\Gamma)}$

and strongly continuous to t in norm $L_2(\Gamma)$.

We introduce the state space of parabolic systems and auxiliary space. For this we consider the bilinear form

$$\ell(\mu, \nu) = \int_{\Gamma} (a(x)\mu'(x)\nu'(x) + b(x)\mu(x)\nu(x)) dx$$

$((\cdot)'$ – the generalized derivative) with fixed measurable and bounded on the Γ_0 functions $a(x)$, $b(x)$, square integrable.

The space $W_2^1(\Gamma)$ includes a set $\Omega_a(\Gamma)$ of functions $u(x) \in C(\Gamma)$, satisfying the relations

$$\sum_{\gamma_j \in R(\xi)} a(1)_{\gamma_j} u'(1)_{\gamma_j} = \sum_{\gamma_j \in r(\xi)} a(0)_{\gamma_j} u'(0)_{\gamma_j} \quad \text{at all nodes}$$

$\xi \in J(\Gamma)$ ($R(\xi)$ and $r(\xi)$ – a plurality of edges, respectively oriented "to the node ξ " and "from the node ξ ",

$u(\cdot)_{\gamma}$ – restriction of the function $u(\cdot)$ on edge γ). Closure

in norm $W_2^1(\Gamma)$ of the set $\Omega_a(\Gamma)$ will be denoted by

$W^1(a, \Gamma)$ (if we assume that the functions $u(x)$ from $\Omega_a(\Gamma)$ are also satisfy the boundary condition $u(x)|_{\partial\Gamma} = 0$, we obtain a space $W_0^1(a, \Gamma)$). Suppose further $\Omega_a(\Gamma_T)$ – a variety of functions $u(x, t) \in V_2(\Gamma_T)$, whose traces are determined on section of the region Γ_T by the plane $t = t_0$ ($t_0 \in [0, T]$) as function of class $W^1(a, \Gamma)$ and satisfy

$$(1) \quad \sum_{\gamma_j \in R(\xi)} a(1)_{\gamma_j} u_x(1, t)_{\gamma_j} = \sum_{\gamma_j \in r(\xi)} a(0)_{\gamma_j} u_x(0, t)_{\gamma_j}$$

for all nodes $\xi \in J(\Gamma)$. The closure of the set $\Omega_a(\Gamma_T)$ in norm (1) let as $V^{1,0}(a, \Gamma_T)$; it's clear that $V^{1,0}(a, \Gamma_T) \subset W_2^{1,0}(\Gamma_T)$. Another subspace of the space $W_2^{1,0}(\Gamma_T)$ is $W^{1,0}(a, \Gamma_T)$ – closure in the norm $W_2^{1,0}(\Gamma_T)$ of the set of smooth functions satisfying (2) for all nodes $\xi \in J(\Gamma)$ and for all $t \in [0, T]$ (space $W^1(a, \Gamma_T)$ is entered similar); $V^{1,0}(a, \Gamma_T)$ – the state space of a parabolic system $W^{1,0}(a, \Gamma_T)$ and $W^1(a, \Gamma_T)$ – auxiliary spaces.

In the space $V^{1,0}(a, \Gamma_T)$ consider initial boundary value problem

$$(2) \quad \begin{aligned} y_t(x, t) - (a(x)y_x(x, t))_x + b(x)y(x, t) = \\ = f(x, t), \end{aligned}$$

$$y|_{t=0} = \psi(x), \quad x \in \Gamma, \quad a(x)y_x|_{x \in \partial\Gamma_T} = \varphi(x, t).$$

Here $(\varphi(x, t))$ is the boundary control of the system (2);

$$f(x, t) \in L_{2,1}(\Gamma_T), \quad \psi(x) \in L_2(\Gamma).$$

Definition 1. A weak solution of the boundary value problem

(2) is the function $y(x, t) \in V^{1,0}(a, \Gamma_T)$, that satisfies the integral identity

$$\begin{aligned} \int_{\Gamma} y(x, t)\eta(x, t)dx - \int_{\Gamma_t} y(x, t)\eta_t(x, t)dxdt + \\ + \ell_t(y, \eta) = \int_{\Gamma} \psi(x)\eta(x, 0)dx + \\ + \int_{\Gamma_t} \varphi(x, t)\eta(x, t)dxdt + \int_{\Gamma_t} f(x, t)\eta(x, t)dxdt \end{aligned}$$

for any $t \in [0, T]$ and any function

$$\eta(x, t) \in W^1(a, \Gamma_T);$$

$\ell_t(y, \eta)$ – bilinear form defined by the relation

$$\begin{aligned} \ell_t(y, \eta) = \int_{\Gamma_t} (a(x)y_x(x, t)\eta_x(x, t) + \\ + b(x)y(x, t)\eta(x, t))dxdt. \end{aligned}$$

Theorem 1. Problem (2) is uniquely solvable in the space of weakly $V^{1,0}(a, \Gamma_T)$.

The state $y(x, t) \in V_2^{1,0}(a, \Gamma_T)$ of the system (3) is determined by a weak solution $y(v)(x, t)$ of the problem (2) ($\varphi(x, t) = v(x, t)$); $L_2(\partial\Gamma_T)$ – the space observation: $Cy(v) = My(v)|_{\partial\Gamma_T}$ ($M : L_2(\partial\Gamma_T) \rightarrow L_2(\partial\Gamma_T)$ – continuous linear operator, here $y(v)|_{\partial\Gamma_T}$ – trace of function

$y(v)$ on the surface $\partial\Gamma_T$); functional $J(v)$ that requires minimization of a convex closed set $U\partial \subset U$ is of the form $J(v) = PMy(v) - z_0 P_{L_2(\partial\Gamma_T)}^2 + (Nv, v)_U$ where $z_0(x, t) \in L_2(\partial\Gamma_T)$ – a predetermined observation.

For the problem (2) define the dual state $\omega(v)(x, t) \in W^1(a, \Gamma_T)$ as a function that satisfies the integral identity

$$\begin{aligned} - \int_{\Gamma_T} \omega(v)(x, t)\zeta(x, t)dxdt + \ell_T(\omega(v), \zeta) = \\ = \int_{\partial\Gamma_T} M^*(My(v)(x, t) - z_0(x, t))\zeta(x, t)dxdt \end{aligned}$$

for all functions $\zeta(x, t) \in W^{1,0}(a, \Gamma_T)$.

Theorem 2. To element $u(x) \in U_\partial$ was optimal control of the system (2) is necessary and sufficient that the following relationship are satisfy to:

$$\begin{aligned} \int_{\Gamma} y(u)(x, t)\eta(x, t)dx - \int_{\Gamma_t} y(u)(x, t)\eta(x, t)dxdt + \\ + \ell_t(y(u), \eta) = \int_{\Gamma_t} \psi(x)\eta(x, 0)dx + \\ + \int_{\partial\Gamma_t} u(x, t)\eta(x, t)dxdt + \int_{\Gamma_t} f(x, t)\eta(x, t)dxdt \end{aligned}$$

for all $t \in [0, T]$ and for all function $\eta(x, t) \in W^1(a, \Gamma_T)$;

$$\begin{aligned} - \int_{\Gamma_T} \omega(u)(x, t)\zeta(x, t)dxdt + \ell_T(\omega(u), \zeta) = \\ = \int_{\partial\Gamma_T} M^*(My(u)(x, t) - z_0(x, t))\zeta(x, t)dxdt \end{aligned}$$

for all functions $\zeta(x, t) \in W^{1,0}(a, \Gamma_T)$;

$$\int_{\partial\Gamma_T} (\omega(u)(x, t) + Nu(x, t))(v(x, t) - u(x, t))dxdt \geq 0$$

for all $v \in U_\partial$. Here $y(u) \in V^{1,0}(a, \Gamma_T)$,

$$\omega(u) \in W^1(a, \Gamma_T) \text{ u } \omega(u)(x, T) = 0, \quad x \in \Gamma.$$

Solved the problem of synthesis of the optimal edge control for $U_\partial = U$.

2.2. Mathematical model with distributed parameters in the netlike domain

The ideas presented above apply to the multivariate case. Consider an open bounded domain Ψ of the Euclidean space R^n that has a netlike structure [3, 4], i.e., $\Psi = (\bigcup_k \Psi_k) \bigcup (\bigcup_l S_l)$, where S_l is a surface that separates adjacent domains Ψ_k , $\partial\Psi$ indicates the boundary of Ψ (initially, the smoothness of $\partial\Psi$ is not important). The locus of conjugation of the adjacent domains Ψ_k will be called the node locus and further denoted by ξ ; it represents the union of surfaces $S_l(\xi)$ whose number coincides with the number of

conjugated domains, that is, $\xi = \bigcup_l S_l(\xi)$. Next you enter the

space $L_2(\Psi)^n$, $L_2(\Psi_T)^n$, and $V_0^{1,0}(\Psi_T)$ similar $L_2(\Gamma)$, $L_2(\Gamma_T)$ and $V_0^{1,0}(a, \Gamma_T)$; $V_0^{1,0}(\Psi_T)$ circuit elements $\phi \in D(\Psi)^n$ in norm

$P\phi P = (P\phi P_{L_2(\Psi)}^2 + P\phi_x P_{L_2(\Psi)}^2)^{1/2}$, satisfies conditions

$$Y|_{S_l^-(\xi)} = Y|_{S_l^+(\xi)}, \quad \sum_l \frac{\partial Y}{\partial n_l^-} \Big|_{S_l^-(\xi)} + \sum_l \frac{\partial Y}{\partial n_l^+} \Big|_{S_l^+(\xi)} = 0$$

(here $D(\Psi_T)^n$ is infinitely differentiable in Ψ_T functions with compact supports in Ψ_T , $\text{div} \phi = 0$; $S_l^-(\xi)$ and $S_l^+(\xi)$ mean the unilateral surfaces for $S_l(\xi)$ defined by the direction of the normals n_l^- and n_l^+ to the surfaces $S_l^-(\xi)$ and $S_l^+(\xi)$, respectively).

For a vector function $Y(x, t) = \{y_1(x, t), y_2(x, t), \dots, y_n(x, t)\} \in V_0^{1,0}(\Psi_T)$ ($x = \{x_1, x_2, \dots, x_n\}$) defined in a domain $\Psi_T = \Psi \times (0, T)$ ($T < \infty$), consider the linearized Navier--Stokes system

$$(3) \quad Y_t - \nu \Delta Y + \nabla p = f, Y(x, 0) = Y_0(x), x \in \Psi, Y|_{\partial \Psi} = 0$$

Theorem 3. *The initial boundary value problem (3) has a unique weak solution in the space $V_0^{1,0}(a, \Psi_T)$. A weak solution of the initial boundary value problem (3) continuously depends on the initial data $f(x, t)$ and $Y_0(x)$ (definition of weak problem solving (3) similar to definition 1 on Γ_T).*

Next, we study two types of optimal control problems that are common in applications, namely, distributed control and starting control (with terminal observations). In the former case, control action appears in the right-hand side of the Navier-Stokes system (i.e., defines the density of external forces); in the latter case, it defines the initial condition of the system at $t = 0$. In both cases, the physical problem is to speed up an incompressible viscous multiphase medium to a given vector velocity field by a given (terminal) time $t = T$.

3. Conclusions

Receive necessary and sufficient conditions for the existence of optimal control, similar to the representation in the theorem 2. The described algorithm is applicable to many optimization problems for differential systems whose states are defined by weak solutions of evolutionary equations on similar networks as in the papers [3 – 5]. Interestingly, other researchers considered alternative approaches to the stability analysis [6 – 8] and stabilization [9 – 11] of the solutions to some applications-relevant classes of complex systems, yet with the same treatment of the optimal control existence conditions.

References

- [1] Volkova A.S., Gnilitzskaya Yu.A., Provotorov V.V., On the solvability of boundary-value problems for parabolic and hyperbolic equations on geometrical graphs, *Automation and Remote Control*, Vol. 75 (2014), No. 2, p. 405-412.
- [2] Provotorov V.V., Boundary control of a parabolic system with distributed parameters on a graph in the class of summable functions, *Automation and Remote Control*, Vol. 76 (2015), No. 2, p. 318-322.
- [3] Provotorov V.V., Eigenfunctions of the Sturm-Liouville problem on a star graph, *Sbornik: Mathematics*, Vol. 199 (2008), No. 10, p. 1523-1545.
- [4] Provotorov V.V., Boundary control of a parabolic system with delay and distributed parameters on the graph, *International Conference "Stability and Control Processes" in Memory of V.I. Zubov (SCP)*, 2015, p. 126-128.
- [5] Провоторов В.В. Оптимальное управление параболической системой с распределенными параметрами на графе (Optimal control of parabolic system with distributed parameters on the graph) *Вестник Санкт-Петербургского университета. Серия 10: Прикладная математика. Информатика. Процессы управления*. 2014. № 3. С. 154-163 (*Vestnik of Saint Petersburg University. Series 10. Applied mathematics. Computer science. Control processes*, 2014, issue 3, p. 109-114 (in Russian)).
- [6] Aleksandrov A.Yu., Zhabko A.P. On the asymptotic stability of solutions of nonlinear systems with delay. *Siberian Mathematical Journal*, 2012, Vol. 53, No. 3, p. 393-403.
- [7] Aleksandrov A., Aleksandrova E., Zhabko A. Asymptotic stability conditions for certain classes of mechanical systems with time delay. *WSEAS Transactions on Systems and Control*, 2014, Vol. 9, p. 388-397.
- [8] Aleksandrov A.Yu., Zhabko A.P. On stability of the solutions of a class of nonlinear delay systems. *Automation and Remote Control*, 2006, Vol. 67, No. 9, p. 1355-1365.
- [9] Veremey E.I., Korchanov V.M. Multiobjective stabilization of the dynamic systems from a certain class. *Avtomatika i Telemekhanika*, 1988, No. 9, p. 126-137.
- [10] Veremey E.I., Sotnikova M.V. Plasma stabilization by prediction with stable linear approximation. *Vestnik of Saint Petersburg University. Series 10. Applied mathematics. Computer science. Control processes*, 2011, issue 1, p. 116-133 (in Russian).
- [11] Karelin V.V. Penalty functions in the control problem of an observation process. *Vestnik of Saint Petersburg University. Series 10. Applied mathematics. Computer science. Control processes*, 2010, issue 4, p. 109-114 (in Russian).

FUNDAMENTALS OF THE KINETIC THEORY OF MULTICOMPONENT EMULSIONS

ОСНОВЫ КИНЕТИЧЕСКОЙ ТЕОРИИ МНОГОКОМПОНЕНТНЫХ ЭМУЛЬСИЙ

Prof., Dr. Tech. Sci. Firsov A.N.

Peter the Great St. Petersburg Polytechnic University – St. Petersburg, Russia

E-mail: anfirs@yandex.ru

Abstract: *The paper proposes a mathematical model for describing the dynamics of multicomponent emulsions, based on ideas and methods of the kinetic theory of gases. The methodological basis of the proposed theory is the ideas and methods of the theory of integral kinetic equations.*

Keywords: MULTICOMPONENT EMULSIONS, KINETIC THEORY, MATHEMATICAL MODEL, INTEGRAL KINETIC EQUATIONS

1. Introduction

In some practically important situations (such as the movement of oil through a well), one has to deal with the problem of the motion of a viscous fluid, inside which there are small inclusions in the form of gas bubbles, droplets of water, solid particles, etc. In this paper, we shall consider the case of emulsions¹. This means that, on the one hand, the entire "mixture" (i.e., the liquid plus inclusions) can be considered as a continuous medium, and on the other hand, that in a small element of the volume of the medium there are "sufficiently many sufficiently small particles". In this connection, it seems quite natural to apply the statistical approach in the above-mentioned problem in the spirit of the kinetic theory of gases. However, the theoretical results known to the author in this direction [1 - 7] are connected with concrete (and rather simple) physical models, which does not allow us to sufficiently cover the problem of the motion of emulsions and give its closed mathematical formulation.

We see the main goal of the work, just in order to give a sufficiently general and precise mathematical formulation of this problem, which allows, in particular, to understand the place and role of some simplifying assumptions. From a methodological point of view, the author followed the basic ideas developed by the Leningrad school of aerodynamics of rarefied gas [8], founded by S. V. Vallander.

2. Statement of the problem of modeling multicomponent emulsions based on a kinetic approach

So, we will consider a viscous liquid, inside which there is a very large number of small "particles" (gas bubbles, drops of other liquids, etc.). The fact that these inclusions are emulsions means, in particular, that the whole mixture (liquid plus inclusions) - from a macroscopic point of view - can be considered as a continuous medium². In this connection, it becomes necessary to distinguish three scales of smallness of distances (and volumes). The first of them corresponds to the concept of a small (elementary, physically infinitesimal) volume of a mixture in the hydrodynamic sense. This is the volume within which, on the one hand, the hydrodynamic and thermodynamic quantities related to the mixture as a continuous medium can be regarded as identical, and on the other hand, in which there are a sufficient number of inclusion particles, so that

the latter can be applied a statistical approach.

The second scale corresponds to the concept of a small (from the hydrodynamic point of view) volume of the main fluid. This volume, generally speaking, is of a lower order than the previous one, and its linear dimensions are of the order of the average distance between the particles.

The third scale corresponds to the dimensions of the inclusion particles themselves. These dimensions will be assumed to be small of a higher order than the average distances between the particles.

As the basic elementary volume, it is natural to take an elementary volume corresponding to the first of the mentioned scales of smallness. Hydrodynamic and thermodynamic quantities characterizing the state of the main fluid will be understood as averaged over such an elementary volume.

Finally, we assume that the inclusion particles have a spherical shape and form a set for which one can use the assumptions commonly used in the definition of the concept of "rarefied gas" (the pairing of collisions, the negligibly small duration of the collision time in comparison with the time of free motion, etc., see, for example, [9]).

The nature of the interaction of particles with each other and with the main fluid requires special consideration. Here we will focus on those aspects of this interaction that are essential for the purposes of this article.

It is known that the gas bubbles differ significantly from the other particles (droplets of liquid, solid particles, etc.), both in terms of their individual properties and the effect on the dynamics of the mixture as a whole. First, this difference is manifested in the fact that the dimensions of gas bubbles can change during their free movement inside the main liquid. This change is due to a change in the temperature and pressure of the main fluid and, as a result, the temperature and pressure of the gas inside the bubbles. With a great degree of certainty, we can assume that at any time the gas inside the bubble is in thermodynamic equilibrium with the surrounding liquid (this means that the temperature and pressure of the gas and liquid are coincide). Thus, the radius r of the bubble is related to the temperature T and pressure p of the main liquid by formula

$$r = \left(\frac{3 m R T}{4 \pi \mu p} \right)^{\frac{1}{3}} \quad (1)$$

Where m - is the mass of gas in the bubble, μ - the molecular weight of the gas, R - the universal gas constant. In the general case, we can assume that r is a given function of m , T , p and μ .

The change in the size of the bubble also results the fact that the mixture (as a continuous medium) cannot be regarded as an incompressible fluid, even if the base liquid is incompressible.

The next feature of bubbles is the possibility of their emergence from "nothing" (for example, as a result of chemical reactions occurring in the main liquid), and as a result of their spontaneous decay. The latter, in the simplest case, can be

¹ The presence of gas bubbles essentially distinguishes the problem of the motion of emulsions from the problem of the motion of suspensions; the latter was quite well developed from different points of view [1 - 7].

² In order to distinguish the emulsion mixture as a continuous medium from the actual liquid, in which all these inclusions are found, we shall call the latter "basic liquid".

considered reliably occurring as a result of reaching a certain limiting bubble radius r_0 .

Interaction of particles with each other is simply their direct collision. In this case, however, a phenomenon such as the fusion of two bubbles or droplets, as well as their crushing are possible. We assume that when two bubbles (or droplets) collide, either their merging is possible, or because of the collision, the number of particles does not change (that is, it remains equal to two), and in the collision of two particles of different types (a bubble with a droplet, etc.), these particles change only the velocities.

3. Basic properties of models and connection with problems of the theory of transfer

We now introduce the basic functions that we shall deal with below, and indicate some of their properties. We will mark the type of particle (bubble, drop, solid particle, etc.) by the indexes i, j, k , etc., taking integral values. The number of values of these indices is obviously finite.

For brevity, we introduce the following terminology. We say that some particle of the type i there is a particle of type (i, x, u, m) , if this particle is reliably located at the point of space with a radius vector $x(x^1, x^2, x^3)$, has a speed $u(u^1, u^2, u^3)$ and mass m , and that this particle is of type (i, x, dx, u, du, m, dm) , if its spatial coordinates are enclosed in the interval $[x^k, x^k + dx^k]$, $k=1,2,3$, the projection of the velocities in the gaps $[u^k, u^k + du^k]$, $k=1,2,3$, and the mass in the gap $[m, m + dm]$.

By the distribution function of the particles of the variety i we shall call the function $f_i(x, u, m, t)$, having the property that the quantity $f_i(x, u, m, t) dx du dm$ gives (up to small higher order) the mathematical expectation of the number of particles of type (i, x, dx, u, du, m, dm) at the moment t . As in [9], it is easy to see that $f_i dx du dm$ there is also a probability of detecting one particle in volume $dx du dm$.

Let us denote by $P_i^{(k)}(x, u, m, t)$ a function possessing the property that the quantity $P_i^{(k)}(x, u, m, t) dt$, $k=2,3,\dots$, is a probability of decay of a particle into k parts over a period of time from t to $t + dt$, if this particle at the time moment t was authentically of the type (i, x, u, m) .

Let's denote by $\omega_i^k(u, m|u_1, m_1)$ the density of mathematical expectation (in the space of variables (u_1, m_1)) of the number of particles of the type $(i, u_1, du_1, m_1, dm_1)$, obtained as a result of authentically decay into k parts of a particle of the type (i, u, m) .

Let us set

$$\hat{T}_i^k(x, u, m, t|u_1, m_1) \equiv P_i^{(k)}(x, u, m, t) \omega_i^k(u, m|u_1, m_1). \quad (2)$$

The following properties of functions ω_i^k и $\hat{T}_i^k$ are obvious:

$$\omega_i^k(u, m|u_1, m_1) = 0 \text{ при } m_1 \geq m, \quad (3)$$

$$\hat{T}_i^k(x, u, m, t|u_1, m_1) = 0 \text{ при } m_1 \geq m, \quad (4)$$

$$\int_0^\infty \omega_i^k(u, m|u_1, m_1) du_1 dm_1 = k, \quad (5)$$

$$\sum_{k=2}^\infty \frac{1}{k} \int_0^\infty \hat{T}_i^k(x, u, m, t|u_1, m_1) du_1 dm_1 = P_i(x, u, m, t) \quad (6)$$

Here

$$P_i(x, u, m, t) = \sum_{k=2}^\infty P_i^{(k)}(x, u, m, t)$$

gives, obviously, the probability of decay per unit time of a particle of the type (i, x, u, m) at the time t . Here and below, if the region of integration is not indicated, we mean the entire space R^3 .

We note that, since the quantity $P_i^{(k)} / P_i$ is the probability of decay of a particle of the type (i, x, u, m) in k parts provided that some decay has occurred reliably, the value

$$\hat{T}_i(x, u, m, t|u_1, m_1) = \frac{1}{P_i} \sum_{k=2}^\infty P_i^{(k)}(x, u, m, t) \omega_i^k(u, m|u_1, m_1) \quad (7)$$

gives the density of mathematical expectation of the number of particles of type $(i, x, u_1, du_1, m_1, dm_1)$, in the space of variables (u_1, m_1) , as a result of a reliable decay of a particle of the type (i, x, u, m) .

By virtue of (3)

$$\hat{T}_i(x, u, m, t|u_1, m_1) = 0 \quad m_1 \geq m. \quad (8)$$

We denote by $\hat{T}_{ij}^k(x, u_1, u_2, m_1, m_2|u, m)$ the density of mathematical expectation (in the space of variables (u, m)) of a number of particles of type (k, x, u, du, m, dm) , obtained as a result of a reliable collision of particles of types (i, x, u_1, m_1) and (j, x, u_2, m_2) . We note that under the assumptions made in §2 $\hat{T}_{ij}^k = 0$ in the following cases:

- 1) $i = j, \quad k \neq i$;
- 2) $i \neq j, \quad k \neq i, j$;
- 3) $i \neq j, \quad k = i, \quad m \neq m_1$;
- 4) $i \neq j, \quad k = j, \quad m \neq m_2$;
- 5) $i = j = k, \quad m > m_1 + m_2$.

Thus, $\hat{T}_{ij}^k$ are nontrivial only for:

- a) $i = j = k$;
- b) $i \neq j, \quad k = i$; c) $i \neq j, \quad k = j$.

In the cases b) and c) an expression for $\hat{T}_{ij}^k$ must contain factors of the form $\delta(m - m_1)$ or $\delta(m - m_2)$, so that in these cases

$$\int_0^\infty \int_0^\infty \hat{T}_{ij}^k(x, u_1, u_2, m_1, m_2|u, m) du dm = 1 \quad (9)$$

Of particular interest is the case a), since here it is necessary to take into account the possibility of merging two particles of one kind. For brevity, we denote $T_{ii}^i \equiv T_i$. Let $h_i(x, u_1, u_2, m_1, m_2)$ is a probability of merging two reliably colliding particles of types (i, x, u_1, m_1) and (i, x, u_2, m_2) , and $\tilde{P}_i(u, u_2, m_1, m_2|u, m)$ is the probability density (in space (u, m)) of the fact that a particle, which as a result of a reliable fusion of these particles, will

have the type (i, x, u, du, m, dm) . We note at once that the expression for $\tilde{P}_i$ should contain a multiplier $\delta(m - (m_1 + m_2))$. It's obvious that

$$\int_0^\infty \int \tilde{P}_i du dm = 1 \quad (10)$$

Let's further $\tilde{T}_i(u_1, u_2, m_1, m_2 | u, m)$ is the density of mathematical expectation (in the space of variables (u, m)) number of particles of type (i, x, u, du, m, dm) , resulting from a reliable collision (without merging) of particles of the type (i, x, u_1, m_1) and (i, x, u_2, m_2) . For $\tilde{T}_i$ (by the assumptions of §3) we have

$$\int_0^\infty \int \tilde{T}_i du dm = 2 \quad (11)$$

In this way,

$$T_i(x, u_1, u_2, m_1, m_2 | u, m) = h_i(x, u_1, u_2, m_1, m_2) \tilde{P}_i(u_1, u_2, m_1, m_2 | u, m) = (1 - h_i) \tilde{T}_i(u_1, u_2, m_1, m_2 | u, m) \quad (12)$$

From (10) and (11) we obtain

$$\int_0^\infty \int T_i du dm = 2 - h_i \quad (13)$$

Suppose, finally, that the function $\Pi_i(x, u, m, t)$ is such that the quantity $\Pi_i dx du dm dt$ is a mathematical expectation of the number of particles of type (i, x, dx, u, du, m, dm) , arising during a period of time from t to $t + dt$ as a result of processes not associated with collisions and particle decays (for example, as a result of chemical reactions in the main liquid).

We note that all the quantities introduced above can depend (as on the parameters) on the macroscopic characteristics of the main liquid at the corresponding point and at the corresponding instant of time.

In concluding this section, we introduce several functions that are important for the future, connected with the free motion of an individual particle of sort i .

Let at the moment t the particle under consideration is of the type (i, x, u, m) . Then the equation of its motion in the main fluid (pressure, velocity, density and temperature are $p(x, t)$, $v(x, t)$, $\rho(x, t)$, $T(x, t)$ respectively) in the presence of acceleration of gravity g , will have the form

$$\dot{x}(\tau) = u(\tau),$$

$$\dot{u}(\tau) = g - \frac{1}{m} \rho V_i(m, p_\tau, T_\tau) + \frac{1}{m} F_{comp}^{(i)}(m, V_i, p_\tau, T_\tau, v_\tau, u(\tau)) \equiv G_i(m, p_\tau, T_\tau, v_\tau, u(\tau)). \quad (14)$$

These equations must be solved under conditions

$$x(\tau)|_{\tau=t} = x, \quad u(\tau)|_{\tau=t} = u. \quad (15)$$

Here we introduce the following notation:

$$\dot{u} \equiv \frac{du}{d\tau}; \quad p_\tau = p(x(\tau), \tau); \quad T_\tau = T(x(\tau), \tau); \quad v_\tau = v(x(\tau), \tau); \quad R_i(x, u, m, t, t+s) = \exp \left[- \int_t^{t+s} Q_i(x_\sigma^{(i)}, u_\sigma^{(i)}, m, \sigma) d\sigma \right] \quad (22)$$

V_i - denotes the volume of the particle under consideration, $F_{comp}^{(i)}$ - denotes the force of resistance to movement of a particle in the main fluid. In the simplest case, for $F_{comp}^{(i)}$ one can use the Stokes formula

$$F_{comp}^{(i)} = 6\pi\eta r a_i (v - u), \quad (16)$$

Where η - denotes dynamic viscosity of the main fluid; r - particle radius; $a_i = 1$, if it is a solid particle, and $a_i = \frac{1}{3} \frac{2\eta + 3\eta'_i}{\eta + \eta'_i}$, if it is a drop or a gas bubble (here η'_i - dynamic viscosity of a liquid (gas) forming a droplet (bubble)).

Further, we denote by

$$x_\tau^{(i)} \equiv \varphi_i(\tau; x, u, m, t), \quad u_\tau^{(i)} \equiv \psi(\tau; x, u, m, t) \quad (17)$$

the solution of problem (14) - (15). Clearly,

$$x_\tau^{(i)} = x, \quad u_\tau^{(i)} = u \quad (18)$$

For each fixed τ functions (18) give a diffeomorphism of the phase space (x, u) into itself. We denote by

$$D_\tau^{(i)} = D_\tau^{(i)}(x, u, m, t) \text{ Jacobian} \quad (19)$$

$$D_\tau^{(i)} = \frac{D(\varphi_i, \psi_i)}{D(x, u)}$$

In particular, an element of the volume of the phase space $dx_\tau^{(i)} du_\tau^{(i)}$ is associated with an element $dx du$ by the relation

$$dx_\tau^{(i)} du_\tau^{(i)} = |D_\tau^{(i)}| dx du \quad (20)$$

We also note that $D_t^{(i)} = 1$.

4. Derivation of integral kinetic equations of the theory of multicomponent emulsions

The principle of the derivation of a system of integral kinetic equations for functions f_i basically does not differ from usual [9, 10]. We, therefore, indicate here only the necessary changes.

4.1. Probability of free movement.

Let $Q_i(x, u, m, t)$ is the probability of collision or decay per unit time of the particle, which at the moment t authentically had the type (i, x, u, m) .

For Q_i we have the expression

$$Q_i(x, u, m, t) = \sum_j \int_0^\infty \int f_j(x, u_3, m_3, t) (r^{(i)} + r_3^{(j)}) |u_3 - u| du_3 dm_3 + P_i(x, u, m, t), \quad (21)$$

where $r^{(i)}$, $r_3^{(j)}$ - denote radii of particles (i, x, u, m) and (j, x, u_3, m_3) respectively. We emphasize that $r^{(i)}$ и $r_3^{(j)}$ depend, generally speaking, on $x, t, m(m_3)$ - see the formula (1).

Let's further $R_i(x, u, m, t, t+s)$ is a probability of a free (without collision and decay) motion during the time from t to $t+s$ of a particle, which at the moment t authentically had the type (i, x, u, m) . Then

$$R_i(x, u, m, t, t+s) = \exp \left[- \int_t^{t+s} Q_i(x_\sigma^{(i)}, u_\sigma^{(i)}, m, \sigma) d\sigma \right] \quad (22)$$

We note that when the expression (21) is substituted into (22), $r^{(i)}$, $r_3^{(j)}$ should be replaced by $r_\sigma^{(i)}$, $r_{3,\sigma}^{(j)}$.

4.2. Birth function.

We denote by $\phi_i(x, u, m, t)$ the birth function of

particles of sort i , that is, a function possessing the property that the quantity

$$dn_0^{(i)} = \phi_i(x, u, m, t) dx du dm dt$$

is a mathematical expectation of the number of particles of type (i, x, dx, u, du, m, dm) , born within a period of time from t to $t+S$. Clearly,

$$dn_0^{(i)} = dn_1^{(i)} + dn_2^{(i)} + dn_3^{(i)},$$

where $dn_1^{(i)}$ is a mathematical expectation of the number of particles of this type, born as a result of collisions, $dn_2^{(i)}$ - born as a result of decay, a $dn_3^{(i)}$ - arising from other causes.

Using the notation of the preceding section, we obtain

$$\begin{aligned} dn_2^{(i)} &= dx du dm dt \int_0^\infty \int f_i(x, u_1, m_1, t) P_i(x, u_1, m_1, t) \hat{T}_i(x, u_1, m_1, t | u, m) du_1 dm_1; \\ dn_3^{(i)} &= \Pi_i(x, u, m, t) dx du dm dt. \end{aligned}$$

The expression for $dn_1^{(i)}$ has the usual form [10]. In this way,

$$\begin{aligned} \phi_i(x, u, m, t) &= \frac{\pi}{2} \sum_{j,k} \int_0^\infty \int_0^\infty |u_1 - u_2| \left(r_1^{(j)} + r_2^{(k)} \right)^2 f_j(x, u_1, m_1, t) f_k(x, u_2, m_2, t) \times \\ &\times T_{jk}^i(x, u_1, u_2, m_1, m_2 | u, m) du_1 du_2 dm_1 dm_2 + \int_0^\infty \int f_i(x, u_1, m_1, t) P_i(x, u_1, m_1, t) \times \\ &\times \hat{T}_i(x, u_1, m_1, t | u, m) du_1 dm_1 + \Pi_i(x, u, m, t) \end{aligned} \quad (23)$$

The system of integral kinetic equations for f_i is output in the standard way [9], and has the form (in the absence of boundaries³)

$$\begin{aligned} f_i(x, u, m, t) &= f_i(x_{t_0}^{(i)}, u_{t_0}^{(i)}, m, t_0) \times \exp \left[- \int_{t_0}^t Q_i(x_q^{(i)}, u_q^{(i)}, m, q) dq \right] \times \\ &\times \left| D_{t_0}^{(i)} \right| + \int_{t_0}^t \phi_i(x_\tau^{(i)}, u_\tau^{(i)}, m, \tau) \times \exp \left[- \int_\tau^t Q_i(x_q^{(i)}, u_q^{(i)}, m, q) dq \right] \left| D_\tau^{(i)} \right| d\tau. \end{aligned} \quad (24)$$

Here $t_0 < t$ - is arbitrary time moment.

A simple technique analogous to that used in [11] and consisting in the passage to the limit $t \rightarrow t_0$, enables us to obtain from (24) the following system of integro-differential equations

$$\frac{\partial f_i}{\partial t} + u \cdot \frac{\partial f_i}{\partial x} + \frac{\partial (G_i f_i)}{\partial u} = \phi_i - f_i Q_i, \quad (25)$$

Where G_i is given by the relation (14) and all functions are taken at the point (x, u, m, t) .

5. Analysis of the relationship between the macroparameters of the mixture and the main liquid

If given $p(x, t)$, $v(x, t)$, $T(x, t)$ then equations (24) together with equalities (21), (23) form a closed system. However, in practice it is difficult to determine the quantities characterizing the main liquid. On the other hand, (in particular, experimentally) values $p_c(x, t)$, $v_c(x, t)$, $T_c(x, t)$, $\rho_c(x, t)$, characterizing the pressure, velocity, temperature, and density of the mixture as a continuous medium, can be found. Therefore, in order to close the system of equations, it is necessary to add the relations connecting

³ In practically interesting problems, we can assume that there are no solid boundaries in the mixture, or on them the functions f_i are given.

f_i , p_c , v_c , T_c , ρ_c , p , v , T , ρ . Let's do it. In the future, it will be more convenient to use the value of the internal energy density instead of the temperature. So, let $E(x, t)$ - is a density (per unit mass) of the internal energy of the main fluid. $E_c(x, t)$ - density (per unit mass) of the internal energy of the mixture (as a continuous medium) and e_i - density (per unit mass) of the internal energy of the substance (gas, liquid) in the particle of the variety i . With the assumptions made in §3, we can assume that $e_i = e_i(p, T)$. Let dx - a certain elementary volume of the mixture (see §3) adjacent to the point x . Obviously,

$$dx = d_1 x + d_2 x,$$

where $d_1 x$ - is a part of the volume dx , occupied by particles, and $d_2 x$ - part of the volume dx , occupied by basic fluid.

We have

$$d_1 x = dx \sum_j \int_0^\infty \int V_j(m, p, v, T) f_j(x, u, m, t) du dm$$

where V_i - is a volume of a particle of a variety i .

The mass of the particles inside dx will be equal

$$d_1 M = dx \sum_i \rho_i \int_0^\infty \int V_i f_i du dm = dx \sum_i \int_0^\infty \int m f_i du dm,$$

where ρ_i - is a density of matter in the corresponding particle.

The mass of the main liquid inside dx is equal, hence,

$$d_2 M = dx \rho \left(1 - \sum_i \int_0^\infty \int V_i f_i du dm \right).$$

Thus, for the density of the mixture we obtain

$$p_c(x, t) = \rho - \sum_i \left(\frac{\rho}{\rho_i} - 1 \right) \int_0^\infty \int m f_i du dm \quad (26)$$

Let us turn to the expression for the macroscopic velocity of the mixture $v_c(x, t)$. We use the formula

$$dK_c = v_c \rho_c dx,$$

where dK_c - is the amount of motion of the mixture inside dx .

Taking into account that dK_c is composed of the amount of motion of the particles and the main fluid, we obtain

$$v_c = \frac{1}{\rho_c} \left[\rho v \left(1 - \sum_i \frac{1}{\rho_i} \int_0^\infty \int m f_i du dm \right) + \sum_i \int_0^\infty \int m u f_i du dm \right] \quad (27)$$

To find the expression for E_c , note that the total energy of the mixture inside dx is made up of the kinetic and internal energies of the particles and the main fluid. Carrying out the corresponding calculations, we obtain

$$E_c = U_c / \rho_c - v_c^2 / 2, \quad (28)$$

Where U_c - is the bulk density of the total energy of the mixture:

$$\begin{aligned} U_c &= E \rho - \sum_i \left(\frac{E \rho}{e_i \rho_i} - 1 \right) \int_0^\infty \int m e_i f_i du dm + \\ &+ \frac{\rho v^2}{2} \left(1 - \sum_i \frac{1}{\rho_i} \int_0^\infty \int m f_i du dm \right) + \sum_i \int_0^\infty \int \frac{m u^2}{2} f_i du dm \end{aligned} \quad (29)$$

Formulas (26) - (29) give the required additional relations for the closure of the system of equations (24).

6. Concluding remarks

Let us now make a few general remarks.

1. Equations (24), (25) are externally similar to the known kinetic equations for gas mixtures [10], but their content is largely different. In particular, under the assumption that there are no collisions between the particles in the problem under consideration,

the right-hand side of (25) does not vanish:

$$\begin{aligned} \phi - f_i Q_i = & \int_0^\infty \int f_i(x, u_1, m_1, t) P_i(x, u_1, m_1, t) \times \\ & \times \hat{T}_i(x, u_1, m_1, t | u, m) du_1 dm_1 + \Pi_i(x, u, m, t) - \\ & - f_i(x, u, m, t) P_i(x, u, m, t) \end{aligned} \quad (30)$$

2. «Equilibrium solution» (that is, a solution that does not depend on x, t) plays a somewhat different role in the problem under study than in the kinetic theory of gases. This is because the dependence on the coordinate is laid inside the very essence of the problem: macro parameters $p(x, t)$, $T(x, t)$ etc. *a priori* are changing over x very quickly.

The latter means that solutions close to equilibrium are of little interest from a practical point of view, and therefore the equilibrium solutions themselves are of little interest. They, however, can be considered as limiting (at $|x| \rightarrow \infty$) function values f_i ; so the equilibrium solutions can be used mainly in the formulation of boundary value problems in unbounded domains.

On the other hand, stationary solutions may be the most interesting from a practical point of view.

3. In the problem of the motion of emulsions, the viscosity of the main liquid plays an important role, and, generally speaking, it can not be neglected. Indeed, if we neglect the term $F_{comp}^{(i)}$ in (14), then any freely moving particle, after a sufficiently long time, will have an arbitrarily high velocity. Consequently, it may turn out that

$$|u|, |x|, t \rightarrow \infty \quad f_i \neq 0,$$

but it is unreasonable.

4. In general case, theoretical search of functions P_i , $\hat{T}_i$, T_i^k etc., defined in §3, is a very difficult task and is currently hardly feasible. In this connection, in our opinion, the role of experimental studies in this direction increases.

We indicate, however, the formula for $P_i(x, u, m, t)$ in the event that the decay of a particle of a variety i occurs if and only if its radius reaches a certain limiting value $r_0^{(i)}$. As already mentioned, we can assume that $r = \theta_i(p(x, t), T(x, t), m)$.

Let

$$m_0^{(i)} = m_0^{(i)}(p, T, r_0^{(i)})$$

is such that

$$r_0^{(i)} = \theta(p, T, m_0^{(i)}).$$

Then

$$P_i(x, u, m, t) = u \cdot \left(\frac{\partial \theta_i}{\partial p} \frac{\partial p}{\partial x} + \frac{\partial \theta_i}{\partial T} \frac{\partial T}{\partial x} \right) \left| \frac{\partial \theta_i}{\partial m} \right|^{-1} \delta(m - m_0^{(i)})$$

5. In the previous sections we did not take into account the possibility of the proper rotation of the inclusion particles. Accounting for this possibility does not cause fundamental difficulties in the sense of deriving the basic equations, but the solution (in particular, numerical) of these equations is significantly complicated because of the increase in the number of independent

variables for functions f_i . On the other hand, in the broad class of practically interesting cases, the assumption of the absence of proper rotation of spherical particles is justified [1, 2].

7. References

1. Fortier A. Mécanique des suspensions. Masson et C-ie Éditeurs, Paris-VI^e, 1967.
2. Reologiya suspenziy. Seriya: Biblioteka Sbornika «Mekhanika»: Sbornik statey. M.: Mir, 1975. [Rheology of suspensions. Series: Library Collection "Mechanics"]. Moscow: Mir, 1975. (In Russian)
3. Bibik Ye.Ye. Reologiya dispersnykh sistem. L.: Izd-vo Leningr. un-ta, 1981. [Bibik E.E. Rheology of disperse systems]. L.: Leningrad University, 1981. (In Russian)
4. Kelbaliyev G. I. i dr. Mekhanika i reologiya neftyanykh dispersnykh sistem. M.: Izd-vo «Maska», 2017. [Kelbaliev G. I. et al., Mechanics and rheology of petroleum disperse systems]. Moscow: "Mask", 2017. (In Russian)
5. Tsibarov V.A. Kineticheskaya model' psevdoozhizhennogo sloya. – Vestn. Leningr. un-ta, 1975, № 13, s. 106-111. [Tsibarov V.A. Kinetic model of a fluidized layer]. – Vestnik of Leningrad University, 1975, No. 13, p. 106 - 111. (In Russian)
6. Pokrovskiy V.N. Statisticheskaya mekhanika razbavlenykh suspenziy. M.: Izd-vo «Nauka», 1978. [Pokrovsky V.N. Statistical mechanics of dilute suspensions]. Moscow: Nauka, 1978. (In Russian)
7. Protod'yakonov I.O., Tsibarov V.A., Chesnokov YU. G. Kineticheskaya teoriya gazovzvesey. L.: Izd-vo Leningr. un-ta, 1985. [Protod'yakonov I.O., Tsibarov V.A., Chesnokov Yu.G. Kinetic theory of gas-weights]. L.: Leningrad University, 1985. (In Russian)
8. Aerodinamika razrezhennykh gazov. Vyp. 1 – 11. Sbornik statey. Pod red. S.V. Vallandera, R.G. Barantseva, L.: Izd-vo Leningr. un-ta, 1963 – 1983. [Vallander S.V., Barantsev R.G. (editors), Aerodynamics of rarefied gases]. Issues 1 – 11, L.: Leningrad University, 1963 - 1983. (In Russian)
9. Vallander S.V. Uravneniya i postanovka zadach v aerodinamike razrezhennogo gaza. – Aerodinamika razrezhennykh gazov. Vyp. 1. L.: Izd-vo Leningr. un-ta, 1963, s. 7-37. [Vallander S.V.. Equations and statement of problems in aerodynamics of a rarefied gas]. - In: Aerodynamics of rarefied gases. Issue. 1. L.: Leningrad University, 1963, p. 7-37. (In Russian)
10. Vallander S.V., Belova A.V. Integral'nyye kineticheskiye uravneniya dlya smesi gazov s vnutrennimi stepenyami svobody. - V kn.: Aerodinamika razrezhennykh gazov. Vyp. 1. L.: izd-vo Leningr. un-ta, 1963, s. 45-52. [Vallander S.V., Belova A.V. Integral kinetic equations for a mixture of gases with internal degrees of freedom]. - In: Aerodynamics of rarefied gases. Issue 1. L.: Leningrad University, 1963, p. 45-52. (In Russian)
11. Filippov B.V. Variant nestatsionarnykh kineticheskikh uravneniy. - V kn.: Aerodinamika razrezhennykh gazov. Vyp. 1. L.: izd-vo Leningr. un-ta, 1963, s. 67-73. [Filippov B.V. Variant of non-stationary kinetic equations]. - In: Aerodynamics of rarefied gases. Issue 1. L.: Leningrad University, 1963, p. 67-73. (In Russian)

CONCEPTUAL CYBERNETIC MODEL OF TEACHING AND LEARNING

Prof. Ing. Veronika Stoffová, CSc.
Faculty of Education – Trnava University in Trnava, Slovakia
NikaStoffova@seznam.cz

Abstract: The article describes conceptual models of learning and teaching as a managed cybernetic system to achieve the knowledge and skills required by the standards. In the article, the learning process is examined as a system for building a knowledge system of a learner actively which is managed and directed by the tutor or teacher to achieve specified learning goal. The information received by the learner has a multimedia nature. The recipient receives the information through various channels (sensors), sight, hearing, smell, taste and touch. The received information arrives in the short-term memory where it is processed - confronted with the previous knowledge in long-term memory of the individual. New information and knowledge can reinforce and extend the recipient's long-term knowledge system if they are consistent logically. If there is a conflict between new information and old knowledge and information - the conflict needs to be solved. The solution can be the clarification of the knowledge system, correction of misconceptions so they are in line with objective reality and relevant knowledge about the subject of learning. It means, learning is not a constant storage of isolated information and information units in memory, but their transformation into active knowledge which is connected to logical structures and forms the knowledge system of the learner. The actual knowledge system enables an individual to solve non-standard problems not only in everyday life but also in science and technology and in various research areas from which the necessary, suitable and useable knowledge (expert) system is built.

Keywords: MODEL OF LEARNING, MODEL OF MULTIMEDIA LEARNING, CYBERNETIC MODEL OF LEARNING, COGNITIVE PEDAGOGICAL AND PSYCHOLOGICAL MODEL OF LEARNING

1. Introduction

Based on the theory of cognitive learning, learning is a complex process of acquiring knowledge and improving the cognitive abilities of a person. The aim of general education is the wisdom – so the learner becomes a wise man to be able to successfully integrate into other people's society, to share the achievements, material and spiritual values of human society, and to be able to contribute to their development, preserve and transfer them.

The success of human society consists of the fact that it can effectively transfer its knowledge to other generations which are able to further develop it. The continuity in the learning of humanity is ensured by education. A wise society creates a stimulating environment for individuals to learn and to ensure their future development. It creates an access to information sources, provides help, motivation and positive patterns. Education is a lifetime process and it can take various forms. It can be organized and institutionalized which means to take place in specially designed institution (e.g. school) or to take place outside the educational institution during everyday life consciously or unconsciously. The institution confirms formal education obtained at a school with a certificate of education. Knowledge, skills and other personal qualities, which society considers to be significant, are informal evidence of the education of an individual. A person, who wants to use the achievements of human society, must be aware of them. Therefore, the acquisition of knowledge is one of the main aims of education. It is obvious that the knowledge of the individual cannot contain everything. A person chooses what he/she wants and needs to learn. His/her choices are influenced by affects, surroundings and life situations in which he/she is involved.

It is necessary to realize that the knowledge is not transferred to a student, but it is created by thinking. Each information and knowledge acquired by a learner must be transformed into a knowledge by the learner – incorporating them into the learner's own knowledge system. In order for education to be successful, the cognitive abilities and thinking of the learners must be developed by the teachers. It depends on the learners to what depth and quality the knowledge will be achieved, in what way an individual will be able to use it in his/her life, and whether he/she will be able to contribute to its creative development.

Teaching in the school environment will be considered as a controlled learning that aims to achieve the prescribed standard of knowledge. Each individual builds his/her knowledge system for his/her whole life. The knowledge system of an individual is not an encyclopaedia of isolated information units, but a system of knowledge linked to relations expressing the context into one

organic whole. In school education, there is a selection of knowledge that a student should master, contained in a curriculum which is created by experts in education in each country. The role of a school is to create a stimulating environment in which the student is motivated to accept this selection of knowledge [14]. School education is the most effective form of education since it focuses on fulfilling the prescribed standards.

2. The Cybernetic Model of Education

It is possible to present the model of the educational system using a control system where the real system is the learner himself, in fact his knowledge system. The process which we manage is student's learning, that means building of the knowledge system of the learner who is supposed to achieve standards in each of the parameters or the prescribed level of the profile of the graduate.

The motivation, the work of the teacher, the acquisition and processing of information (a process of learning) changes the real object. The ideal system which is compared to the real system is expressed by the standards. In case that the learner does not reach the prescribed level of knowledge, based on the difference in knowledge, it is necessary to apply the feedback regulator represented by the teacher, by the influence of the environment and the surroundings, by acquiring new knowledge from different sources, that means learner's process of learning. The feedback regulator does not respond to the positive difference, the situation when the learner knows more than it is prescribed by the standards. However, this may have a positive effect on the activities of the learner and may increase the motivation to self-study [13], [14].

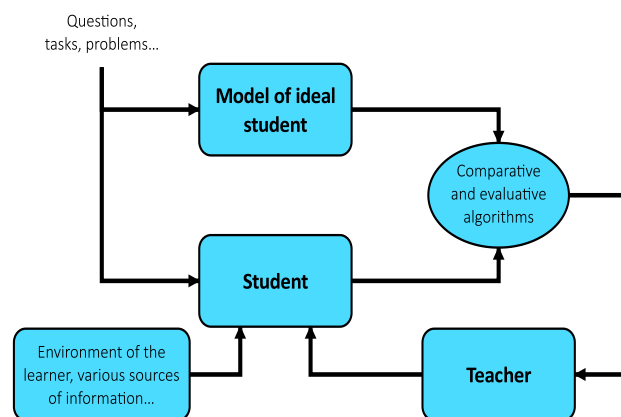


Fig. 1 Cybernetic model of school education

3. The model of multimedia teaching and learning

Learner perceives information and knowledge by several senses. Information is mostly gathered by **sight - eyesight**. According to Fredmann, almost 83% of information we sense with eyesight. Even our research, as well as other researches approved this hypothesis. Graphical information in form of a picture, an animation, a graph etc. can be considered as concentrated form of presenting particular knowledge. Hearing is second mostly used sense for gathering information with 11%. Other senses, such as touch or taste are not as important as sight and hearing. Researches show that we are able to remember only 20% of what we have heard or read, 30% of what we have seen. We are able to remember almost 70% of all information by integration of previous activities with feedback. To be able to remember 80% and more information it is necessary to effectively use gathered information and transfer it into personal system of knowledge. This will help with deduction of new knowledge [03], [04], [16].

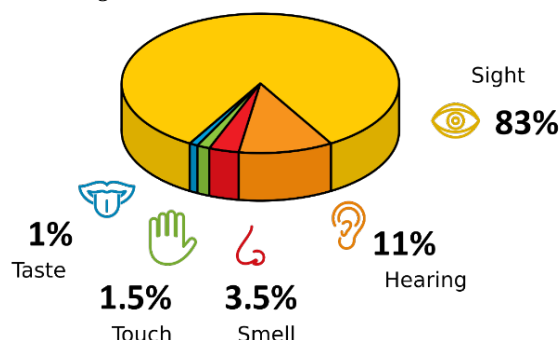


Fig. 2 How do we get information and knowledge

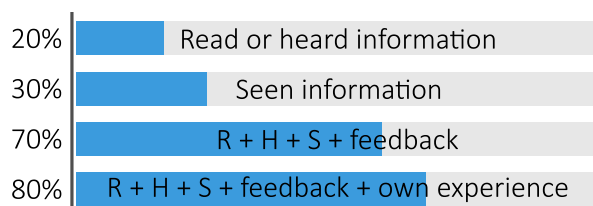


Fig. 3 How much we remember from the acquired knowledge

This theory and experimental results are in accordance with Niemierko's and Bloom's taxonomies of educational goals [09], [13].

General taxonomy of cognitive objectives describes 7 levels of how to handle a topic - gathering knowledge. 1. Knowledge, 2. Comprehension, 3. Application, 4. Analysis, 5. Synthesis, 6. Evaluation, 7. Imagination.

Niemierko's taxonomy of cognitive educational goals consists of four levels and is used especially for technical and exact sciences. **Level of knowledge:** 1. remembering; 2. understanding. **Level of competences:** 3. Specific transfer - application of acquired knowledge according presented tasks, in standard situations; 4. Non-specific transfer - an application in new and non-specific - real problem situations.

Niemierko's taxonomy of cognitive aims can be transformed into the knowledge requirements of the learner. The learner knows:

- 1) To list, to reproduce, to repeat, to name;
- 2) To explain, to describe, to give examples, to say/express with his/her own words;
- 3) To apply, to calculate, to demonstrate, to quantify;
- 4) To compare, to judge, to defend, to justify, to draw conclusions, to generalize, etc.

Bloom's Taxonomy (created in 1956) in order to promote higher forms of thinking in education, such as analysing and evaluating concept, process, procedures and principles, rather than just remembering facts (rote learning) [02]. It contains 6 linear consecutive cognitive levels: Evaluation, Synthesis, Analysis, Application, Comprehension and Knowledge. The cognitive domain involves knowledge and the development of intellectual skills. This

includes the recall or recognition of specific facts, procedural patterns, and concepts that serve in the development of intellectual abilities and skills. This six major categories of cognitive processes, starting from the simplest evaluation to the most complex knowledge (see the Fig. 4) for an in-depth coverage of each category. The categories can be thought of as degrees of difficulties. That is, the first ones must normally be mastered before the next one can take place.

The original Bloom's taxonomy was revised after 45 years [01]. The order of synthesis and evaluation categories was exchanged. The synthesis has been replaced by the "create" category, which is not understood as a rebuilding of individual elements, but includes a creative element along with the evaluation (the usage of critical thinking). The category of comprehension was renamed to understanding. The cognitive dimensions are expressed in the verb form, while the knowledge dimensions are in the form of a noun (substantive). Bloom's Taxonomy is mostly used when designing or modelling educational, training, and learning processes

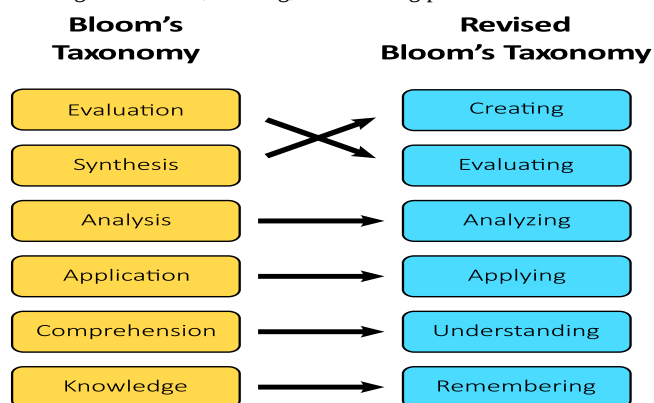


Fig. 4 Bloom's original and revised taxonomy

The chart (Fig. 4) compares the original taxonomy with the revised one changing the names in the six categories from noun to verb forms.

Revised Bloom's taxonomy recedes from the hierarchical order of the categories from the lowest to the highest. Overlapping may occur mainly during the educational activities, e.g. the learner sometimes evaluates even though he/she had not analysed the subject matter, or he/she creates new knowledge on the basis of partial knowledge. However, all dimensions should be shown. Table 1 provides examples of possible student activities for each of the component of the knowledge and cognitive dimension. Insertion the objective into the table is not always so easy and unequivocal. The formulation of the educational objectives and their taxonomic classification helps the teacher to choose the teaching methods (methods and forms of teaching) as well as it helps him/her in assessment, which is appropriate to the nature of the objectives [09]. The activity of the teacher and the learner should correspond to the chosen educational objective. Today, Bloom's taxonomy is easily understood and probably the most often applied one.

Table 1: examples of possible student activities for each of the component of the knowledge and cognitive dimension

The Knowledge Dimension	Remembering	Understanding	Applying	Analyzing	Evaluating	Creating
Factual Knowledge	List	Summarize	Sort, classify	Order	Choose	Combine
Conceptual Knowledge	Describe	Interpret, recognize	Experiment	Explain, compare	Estimate, determine	Plan, outline
Procedural Knowledge	Organize	Predict	Calculate, solve	Differentiate, depict	Conclude	Compose, design
Meta-Cognitive Knowledge	Appropriate Use	Process	Construct	Create	Perform, express	Actualize, improve

The effectiveness of learning is closely related to the participants' activity. Already, Edgar Dale (1900-1985), an American educational researcher, has found this in conjunction with developing the "Cone of Experiences" [05], which shows the effectiveness of the different teaching methods. From this we can see that the best methods are based on being more reliant on learner activity (Fig. 5). For this reason, today's most popular educational methods are all focused on activities of learners such as Inquiry-Based Learning, Problem-Based learning or Project-Based Learning. At the top of the pyramid are the less-active methods, where we listen and read the learning materials. In order for the educational environment to be effective, action, interaction and creativity should be introduced into education. This methods are located at the bottom of the cone [03], [04], [05], [06], [15].

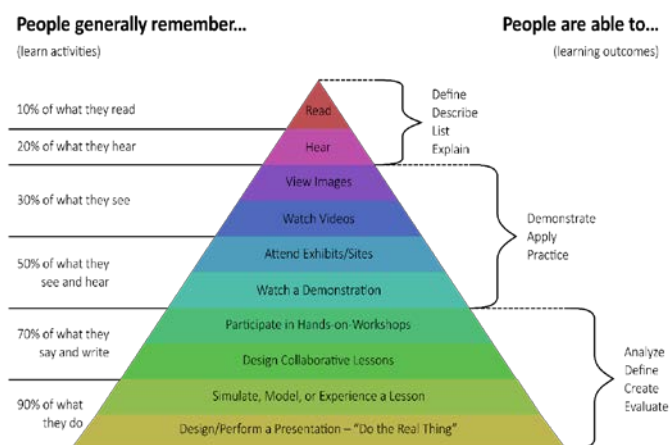


Fig. 5 Cone of Experiences

4. The Model of Knowledge System Creation

The teacher in the educational process affects the learner with the intention of achieving educational goals. Teaching is considered to be an optimal way of getting necessary knowledge. In implementing the learning model, it is important to apply both the taxonomy of objectives as well as the effectiveness of the various sources of knowledge and information that are acquired through different channels and the existence of different learning methods and styles. An important part of the learning model is the **presentation of knowledge** and information which are the subject of learning. This subsystem is part of the **model of the (ideal) teacher**. Learning material should be presented with the appropriate pedagogical transformation, in a multimedia structured form, with the respect to the student's mental level and his/her active involvement in the learning process. Another model of the subsystem of the teacher is **the management of the learning process**. It is necessary to identify learners, their learning style, their current level of knowledge, etc. to optimally manage the process of learning in order to individualize the learning process and shorten the time that is necessary to process gained information and knowledge. Incorporating new knowledge into the learning system of the learner means confronting new knowledge with knowledge in one's own knowledge system. In the case that the new knowledge is in accordance with the existing ones, the knowledge is deepened, and its durability is increased [16], [19]. If a conflict arises in assessing and processing new information and knowledge with the old ones, it needs to be solved. Thus, the individual constantly builds and improves their own knowledge system not only by consciously searching for new information and by checking and using it in practice, but also by the influence of the environment where the learner lives and in contact and communication with others. The objective of building the knowledge system is to make the individual wise, to change the quantity into the quality, so that understood (passive) information and knowledge become active, and therefore the individual is convinced of their relevance.

The knowledge system of an individual is not an isolated information and knowledge, but a system of knowledge – knowledge which is interconnected by relations that express the connection between knowledge. The knowledge system contains knowledge that is useful for solving everyday life problems, problems in professional life, employment problems and the problems of society we live in.

The learner analyses the acquired information, confronts them with knowledge of his/her own knowledge system, tests them by active usage, and incorporates them into his/her own knowledge system by synthesis, in other words the new information become knowledge. (See the Fig. 5.) It is according the rule of learning where the ultimate goal is a wise learner. „From information and knowledge to wisdom“.

In the Fig. 5 can be seen how the learner processes information and knowledge that is presented in different forms and how he/she creates his/her own knowledge system.

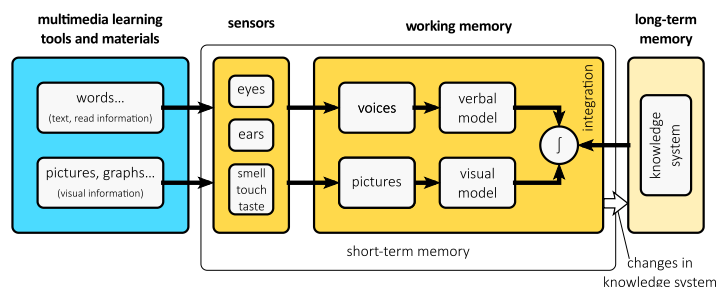


Fig. 6 Model of knowledge system creation

5. The Knowledge System Building

In constructing the graphical model of building the knowledge system, we proceeded from the cognitive theory of multimedia learning of Mayer. Mayer, based on his research, developed the Cognitive Theory of Multimedia Learning, based on 12 principles, which should be taken into account when developing multimedia teaching materials. These principles were grouped into three main categories [08].

- I. Reducing processing of information not covered (Coherence Principle; Signalling Principle; Redundancy Principle; Spatial Contiguity Principle; Temporal Contiguity Principle).
- II. Control the processing of relevant information (Segmenting Principle; Pre-training Principle; Modality Principle).
- III. To promote constructive processing of information (Multimedia Principle; Personalization Principle; Voice Principle; Image Principle).

In the Fig. 5 is illustrated how a knowledge system is built and it is also easy to see how the human brain processes several types of information.

On the basis of the results of Mayer's research, the basic principles of multimedia learning have been developed. Those results are confirmed by our research results in the area of the usage of the multimedia learning aids for programming in forms of animation-simulation models, which we have accomplished during the past years at J. Selye University in Komárno as well as at Trnava University in Trnava. The animations were developed in HTML5 using JavaScript technologies. We also used the CreateJS libraries (www.createjs.com) for animating the objects.

Our research results have shown that when the explanation of teaching material is in textual form, it is important to give students enough time to read and comprehend it. According to the principles of Mayer's multimedia learning, it is not recommended to show the text and animate the objects on the screen at the same time. It is better to let students start the animation themselves after reading and understanding the explanation. Another possible solution could be to use narration during the animation instead of textual explanation [16], [17], [18], [19].

The results of several experiments emphasized that the animations may be used more efficiently when they are not displayed alone, but when they are part of a learning environment. This environment could be an electronic textbook with hypertext structure, enriched with embedded animations, diagrams, and examples [19].

The limited scope of the article does not allow detailed analysis of the individual subsystems of the model. We plan to publish this analysis in the next article about the modelling of the learning process.

6. Conclusion

The model of learning – building of the knowledge system of the learner can be implemented in various ways. It might be built as a learning expert system [12], a virtual “second life” university or as a learning environment created with the support of LMS and CMS, where the presentation of knowledge can have a different form, e.g. multimedia interactive animation-simulation model of the studied phenomenon, or dynamic process [10], [11]. The management and optimisation of the learning process should be realised on the basis of the logical hyper-structure of the knowledge of the studied object, which is the subject of the education. In that case the logical hyper-structure of the knowledge contributes to the systematisation of the knowledge and to the correct improvement and building of the knowledge system of the learner. These kinds of education systems must be implemented in a such way that their utilization reduces the time needed to understand new knowledge, to process them and to convert them into active knowledge, that are useful and easily accessible in the long-term memory of the learner. They help him/her solve various problems, not only in the professional field, but also in everyday life. If we want educational software systems, which support learning, to deliver the expected effect, we must respect the rules of creating and building a knowledge system and the pedagogical, psychological and didactic rules of learning.

References

- [01] Anderson, L. et al. (2001) A Taxonomy for Learning, Teaching a Assessing of Educational Objectives. New York: Longman, 2001. 352 s. ISBN 0-321-08405-5.
- [02] Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *The Taxonomy of Educational Objectives, The Classification of Educational Goals, Handbook I: Cognitive Domain* (B. S. Bloom Ed.). New York: David McKay Company, Inc.
- [03] Czakoóv, K. (2016) Creation small educational software in the micro-world of small languages. In: *Teaching Mathematics and Computer Science*. 14th volume, issue one, 2016/1, p. 117. Debrecen : University of Debrecen, 2016. ISSN 1589-7389
- [04] Czakoóv K. (2016) Felfedezsen alapul aktiv tanuls mikrovilg krnyezetben. (Discovery-based active learning in microworld environment). In: *New methods and technologies in education and practice : 19. DIDMATTECH 2016*. Budapest : Etvs Lornd University, 2016. 168-173. s. ISBN 978-963-284-799-3
- [05] Cone of learning: (On-line <http://bit.ly/1TNR2o6>, last access: 2017. 11.10)
- [06] Ills, Z., H. Bakonyi V. (2016) Experiment for increasing equal opportunity in university with the support of a BYOD system, SzmOkt 2016, Kolozsvr október 8-9. ISSN 1842-4546
- [07] Ills, Z. (2017) Vals idej rendszerek s megjelensk az oktatsban - Real-time systems and their appearance in education. (Habilitation theses). ELTE- FI, Budapest 45. p. 2017
- [08] Mayer, R. E. (2009). *Multimedia Learning* (second ed.). New York, USA: Cambridge University Press.
- [09] Majherov, J.: Revidovan Bloomov taxonoma a kompetencie pre pouzvane IKT - Revised Bloom's Taxonomy and Competencies for Use of ICT, on-line access - https://www.pdf.umb.sk/~lrovnanova/taxonomia_ciele_Anderson.pdf
- [10] Stoffa. V. (2004) Modelling and simulation as a recognising method in the education. *Educational Media International*, 41(1):51-58, 2004.
- [11] Stoffa V. (2008) Az animci szerepe az elektronikus tanknyvekben. *Informcis trsadalom*, VIII(3):113-125, 2008.
- [12] Stoffov, V., Gergelov, Z. (2001) Expert systems in the systematisation of the knowledge at primary schools (Expertn systmy v systematizcii poznatkov na Z. *Technolgia vzdelvania : Zvzok 3 : Educational Technology : Volume 3*. 1. vyd. Nitra, Pedagogick fakulta UKF v Nitre, 2001, s. 160-165.
- [13] Stoffov, V. (2004) *Pota – univerzlny didaktick prostriedok* 1. vyd. Nitra : Fakulta prrodnch ved UKF v Nitre, 2004. 172 s. ISBN 80-8050-450-4
- [14] Stoffov, V. (2006) The Importance of Didactic Computer Games in the Acquisition of New Knowledge. In: *The European Proceedings of Social & Behavioural Sciences EpSBS*. pp. 676-688. eISSN: 2357-1330. (on-line access: <http://dx.doi.org/10.15405/epsbs.2016.11.70>)
- [15] Stoffov, V. - Czakoóv, K. (2016) Prostredie na uenie sa bdanm. In: *vod do programovania v prostred mikrosvetov : vysokoskolsk uebnica*. Komrno : Univerzita J. Selyeho, 2016. 8-33. s. ISBN 978-80-8122-170-5
- [16] Vgh, L. – Stoffov, V. (2016) An interactive animation for learning sorting algorithms: How students reduced the number of comparisons in a sorting algorithm by playing a didactic game. In: *Teaching Mathematics and Computer Science*. Debrecen : Institute of Mathematics – University of Debrecen, 14th volume, issue one, 2016/1, s. 45–62. ISSN 1589-7389.
- [17] Vgh, L. (2011) Animations in teaching algorithms and programming (Animcie vo vyuovan algoritmov a programovania). In Ji Dostl, editor, *Nov technologie ve vzdelvan*, pages 47-51, Olomouc, CZ, 2011. Palack University, Olomouc.
- [18] Vgh, L. (2011) From bubblesort to quicksort with playing a game (Hravou formou od bublinkovho triedenia po rchle triedenie). In Ji Neubauer and Eva Hjkov, editors, XXIX. International Colloquium on the Management of Educational Process, pages 539-549, Brno, CZ, 2011. University of Defence.
- [19] Vgh, L. (2017) A programozs tanulsnak s tantsnak tmogatsa elektronikus tananyagba bepthet interaktiv animcis modellekkel. (PhD theses) ELTE - Faculty of Informatics, Budapest 202. p. 2017

MODELING AND SIMULATION OF INDUSTRIAL PROCESSES

Christo Boyadjiev

Institute of Chemical Engineering, Bulgarian Academy of Sciences
chr.boyadjiev@gmail.com

Abstract: In the paper are presented the main theoretical techniques used for the modeling and simulation of industrial processes. The main focus is on the physical side of the theoretical techniques and their mathematical side is reduced to a reasonable minimum. Different theoretical approximations as thermodynamic and hydrodynamic levels are used.

KEYWORDS: MODELING, SIMULATION, THEORETICAL APPROXIMATION, THERMODYNAMIC LEVEL, HYDRODYNAMIC LEVEL, MECHANISM IDENTIFICATION, PARAMETERS IDENTIFICATION, STATISTICAL ANALYSIS.

1. Introduction

The modeling and simulation are a basic approach to quantifying processes and phenomena [1-3]. They have become realistic as a result of the development of computing and applied mathematics. In the industry, the modeling and simulation offer a quantitative description of the kinetics of processes and systems for the purposes of their optimal design or control. In the industrial systems, the process models of the individual devices are known in advance, and system models offer quantitative descriptions of the systems (process systems engineering).

In the industry, quantification of systems can also be used for various other tasks. For example, in periodically operating systems, the optimal schedules of the apparatuses (machines) for conducting different processes (operations) of different duration and different sequence to obtain different substances (machined details) can be determined. In these cases, the mass service theory offers models that allow for optimal solutions.

The fundamentals of the modeling and simulation, as a part of human knowledge and science, are related to the combination of intuition and logic. They are in different scales in the individual sciences [4, 5]. In the mathematics, logic dominates intuition, where intuition is the axiom (unconditional truth that is not proof able), and logic is the theorem (the logical consequences of the axiom). In the natural sciences (physics, chemistry, biology), the logic/intuition ratio is maintained, but axioms are usually conditional (principles, postulates, laws). This ratio goes back to the humanities and reaches the extreme in religion.

The modeling and simulation offer quantitative (mathematical) descriptions that have different degrees of detail. The lowest level is the thermodynamic (non-equilibrium thermodynamics) that examines the volume of the phase (gas, liquid, solid). The next level is the hydrodynamic, which examines the elementary phase volumes (mechanics of continua), which are much smaller than the phase volumes, but much larger than the intermolecular volumes, i.e. the molecules are indistinguishable. The highest level is the molecular (the kinetic theory of the ideal gas).

The modeling and simulation of industrial processes has a wide application, so the processes in the chemical industry and related biotechnologies and heating technologies, will be discussed. The major part of these processes is the transfer of mass and heat as a result of phase or phase boundary reactions. By reaction, will be understand the creation or disappearance of a particular substance (or amount of heat) as a result of a chemical reaction in the phase or on the phase boundary, interphase mass transfer, adsorption on the phase boundary or a liquid-vapor-liquid phase transition. These reactions result in varying concentrations and temperatures in the phases, i. e. to a deviation from the thermodynamic equilibrium and as a result of the mass transfer and heat transfer to restore the thermodynamic equilibrium. The models of the mass transfer and heat transfer are analogous and, therefore, the models of mass transfer in industrial processes will be presented.

2. Thermodynamic approximation

The reactions deviate the industrial systems from the thermodynamic equilibrium and the industrial processes for its recovery begin. The determining of the rate of these processes is a major problem in the industry, as it is the basis for their optimal design and control. This gives reason to use the thermodynamic laws of irreversible processes such as mathematical structures in the construction of the process models, described by extensive and

intense variables (in the case of merging of two identical systems, the extensive variables double their values, while the intensive variables retain their values).

The kinetics of the irreversible processes uses the mathematical structures, resulting from Onsanger's "linearity principle" [6]. According to him, the mean values of the derivatives at the time of the extensive variables depend linearly on the mean deviations of the conjugated intensive variables from their equilibrium states (values). This principle is valid in the vicinity of the equilibrium, and proportionality coefficients are kinetic (rate) constants.

According to the principle of linearity, the mass derivative at time

$J_0 = \frac{dm}{dt}$ [kg-mol.s⁻¹] depends linearly on the deviation from the

thermodynamic equilibrium ΔC [kg-mol.m⁻³] of the concentration in two phase volumes or in one phase and the phase boundary, i.e.

$$J_0 = k_0 \Delta C, \quad (1)$$

where k_0 [m³.s⁻¹] is a proportionality coefficient.

Consider a system that contains two identical volumes in one phase $V_1 = V_2 = V$ [m³]. The system contains a substance whose masses

m_i [kg-mol] and concentrations $C_i = \frac{m_i}{V_i}$ [kg-mol.m⁻³] are

different in two volumes, $i = 1, 2, \dots$. The system is not in

thermodynamic equilibrium. Let us assume for certainty

$C_1 - C_2 > 0$, $i = 1, 2$. As a result, the mass of the substance

starts to be transferred from volume V_1 to volume V_2 for to

achieve the equilibrium. According to the principle of linearity, the

mass transfer rate between the two volumes J_0 [kg-mol.s⁻¹] can be represented as:

$$J_0 = \frac{dm_1}{dt} = -\frac{dm_2}{dt} = k_0 (C_1 - C_2), \quad (2)$$

where k_0 [kg-mol⁻¹.m³.s⁻¹] is a proportionality coefficient. If we

replace masses with concentrations $m_i = V_i C_i$, $i = 1, 2$, the mass

transfer rate in one phase J [kg-mol.m⁻³.s⁻¹] between two points with different concentrations is:

$$J = \frac{dC_1}{dt} = -\frac{dC_2}{dt} = k (C_1 - C_2), \quad (3)$$

where k [s⁻¹] is a rate coefficient. This equation is capable of presenting the rate of interphase mass transfer in the case of adsorption or catalytic process, where C_1 is the concentration of the

substance in the gas phase, while C_2 is the concentration of the substance in the gaseous portion of the solid (capillaries of the adsorbent or catalyst) phase.

In the cases, where the volumes $V_1 = V_2 = V$ are in different phases (for example, 1 is a gas phase and 2 is a liquid phase), the

thermodynamic equilibrium law has the form $C_1 - \chi C_2 = 0$, i.e. this is the Henry's law and χ is the Henry's number. If $C_1 - \chi C_2 > 0$ the mass transfer is from phase 1 to phase 2 and the mass transfer rate between phases is:

$$J = k(C_1 - \chi C_2), \quad (4)$$

where k [s^{-1}] is the rate coefficient of the interphase mass transfer. On the surface between two phases, the thermodynamic equilibrium is immediately established, practically, i.e. $C_1^* - \chi C_2^* = 0$, where C_i^* , $i = 1, 2$, are the equilibrium concentrations on the phase boundary. Thus, the mass transfer rate can be expressed by mass transfer rate in two phases:

$$J = k_1(C_1 - C_1^*) = k_2(C_2^* - C_2), \quad (5)$$

where k_i , $i = 1, 2$ [s^{-1}] are mass transfer rate coefficients.

The Onsanger principle of linearity represents the thermodynamic approximation of the mathematical description of the kinetics of irreversible processes, but it does not show the way to reach equilibrium, i.e. the mechanism of the process and as a result the rate coefficient is not known. Obviously, this "thermodynamic level" does not allow a real quantitative description of the kinetics of irreversible processes in industry and the next level of detail of the description, the so-called "hydrodynamic level", should be used.

3. Hydrodynamic approximation

The processes in the chemical industry and related biotechnologies and heating technologies are realized in one-, two- and three-phase systems (gas-liquid-solid). They are a result from the reactions, i.e. processes of disappearance or creation of any substance. The reactions are associated with a particular phase and can be homogeneous (occurring in volume of the phase) or heterogeneous (occurring at the interface with another phase). Homogeneous reactions are usually chemical, while heterogeneous reactions may be chemical, catalytic and adsorption. Heterogeneous reaction is the interphase mass transfer too, where on the interphase boundary the substance disappears (created) in one phase and creates (disappears) in the other phase.

The volume reactions lead to different concentrations of the reagents in the phase volumes and as a result two mass transfer processes are realized – convective transfer (caused by the movement of the phases) and diffusion transfer (caused by the concentration gradients in the phases). The mass transfer models are a mass balance in the phases, where components are convective transfer, diffusion transfer and volume reactions (volume mass sources or sinks). The surface reactions participate as mass sources or sinks in the boundary conditions of the model equations. The models of this complex process are possible to be created on the basis of the mass transfer theory, whose models are created by the models of the hydrodynamics, diffusion and reaction kinetics. The mass transfer theory combines the chemistry, physics and mathematics and builds its logical structures on three main "axioms":

1. The postulate of Stokes for the linear relationship between the stress and deformation rate, which is the basis of the Newtonian fluid dynamics models;
2. The first law of Fick for the linear relationship between the mass flow and the concentration gradient, which is the basis of the linear theory of the mass transfer;
3. The first law of Fourier for the linear relationship between the heat flux and the temperature gradient, which is the basis of the linear theories of the heat transfer.

These are the laws of the impulse, mass and energy transfer.

In Boltzmann's kinetic theory of the ideal gas, these axioms are replaced by the "elastic shock" axiom (in a shock between two molecules the direction and the velocity of the movement change, but the sum of their kinetic energies is retained, i.e. there is no loss of kinetic energy) and the rate coefficients are theoretically

determined by the average velocity and the average free run of the molecules.

The contemporary mass transfer theory is based of diffusion boundary layer theory (Landau, Levich [7]). This approach substitutes (physically justified) elliptic partial differential equations with parabolic partial differential equations, which facilitates their mathematical solution and offers a mathematical description of physical processes with free (not predetermined) ends.

The diffusion boundary layer theory is developed in the cases of drops and bubbles (Levich, Krylov [8]), film flows (Levich, Krylov, Boyadjiev, Beshkov[9, 10]), non-linear mass transfer and hydrodynamic stability (Krylov, Boyadjiev, Babak [11, 12]).

3.1. Mass transfer theory

The complex industrial processes are a collection of elementary physical and chemical processes. For example, the chemical absorption in a packed bed column represents a physical absorption of a gas phase component in the liquid phase and a subsequent chemical reaction with a component of the liquid phase. The gas moves in the column like jets and bubbles, while the liquid moves in the form of drops, jets, and flowing films on the surface of the packed bed. As a result, the chemical absorption in a packed bed column is a combination of many elementary physical and chemical processes, as absorption in the systems gas-liquid drops, liquid-gas bubbles, gas-liquid film flow, etc. As an example will be considered the gas absorption in liquid film with free interface.

Let us consider absorption of a slightly soluble gas in a laminar liquid film [9, 10] in a coordinate system (x, y) , flowing over a flat vertical interface $(y = 0)$. The hydrodynamic model has the form:

$$\nu \frac{\partial^2 u_x}{\partial y^2} + g = 0, \quad \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} = 0; \quad (6)$$

$$y = 0, \quad u_x = 0, \quad u_y = 0; \quad y = h_0, \quad \frac{\partial u_x}{\partial y} = 0$$

and the velocity distribution is:

$$u_x = \frac{g}{2\nu}(2h_0y - y^2), \quad u_y \equiv 0. \quad (7)$$

If these conditions the convection-diffusion model has the form:

$$\begin{aligned} \frac{g}{2\nu}(2h_0y - y^2) \frac{\partial c}{\partial x} &= D \left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2} \right), \\ x = 0, \quad c &= c_0; \quad x \rightarrow \infty, \quad c = c^*; \\ y = 0, \quad \frac{\partial c}{\partial y} &= 0; \quad y = h_0, \quad c = c^*, \end{aligned} \quad (8)$$

where the thermodynamically equilibrium exists at the film interface $(y = h_0)$ and c^* denotes the equilibrium concentration.

The solid interface $(y = 0)$ is impenetrable for the diffusing substance with inlet concentration $C_0 < C^*$ (absorption). A film of length l will be considered.

The diffusion boundary layer thickness δ [2] is:

$$\delta = \sqrt{\frac{Dl}{u^*}} = \frac{l}{\sqrt{Pe}}, \quad Pe = \frac{u^*l}{D}, \quad \frac{\delta}{l} = \frac{1}{\sqrt{Pe}}, \quad u^* = \frac{gh_0^2}{2\nu}, \quad \frac{\delta^2}{h_0^2} = Fo = \frac{Dl}{u^*h_0^2}, \quad (9)$$

where u^* , Pe , Fo , are the film interface velocity, the Peclet and Fourier numbers.

The diffusion boundary layer thickness δ is less than liquid film thickness h_0 that permits the diffusion boundary layer

approximation to be applied. As a result, the next generalized variables can be introduced:

$$x = lX, \quad y = h_0 - \delta Y, \quad c = c_0 + (c^* - c_0)C, \quad (10)$$

$$\text{where } 0 = \frac{h_0}{l} \leq 10^{-2}.$$

The introduction of (10) into (8) yields:

$$\begin{aligned} (1 + FoY^2) \frac{\partial C}{\partial X} &= \frac{Dl}{u^* \delta^2} \left(Pe^{-1} \frac{\partial^2 C}{\partial X^2} + \frac{\partial^2 C}{\partial Y^2} \right); \\ X = 0, \quad C = 0; \quad X \rightarrow \infty, \quad C = 1; \\ Y \rightarrow \infty, \quad C = 0; \quad Y = 0, \quad C = 1, \end{aligned} \quad (11)$$

where Fo is a small parameter:

$$\frac{\delta^2}{h_0^2} = Fo < 10^{-1}, \quad \frac{\delta^2}{l^2} = Pe^{-1} \leq 10^{-2}. \quad (12)$$

The problem (11) in the diffusion boundary layer approximation ($10^{-2} \geq Pe^{-1} = 0$) has the form: namely:

$$\begin{aligned} (1 + FoY^2) \frac{\partial C}{\partial X} &= \frac{\partial^2 C}{\partial Y^2}; \\ X = 0, \quad C = 0; \quad Y = 0, \quad C = 1; \quad Y \rightarrow \infty, \quad C = 0. \end{aligned} \quad (13)$$

The mass transfer rate (J) in the liquid film flow with a length l is the average value of the local mass flux through the face interphase ($y = h_0$). On the other hand this rate can be presented using the mass transfer coefficient k . As a result

$$J = \frac{D}{l} \int_0^l \left(\frac{\partial c}{\partial y} \right)_{y=h_0} dx = k(c^* - c_0). \quad (14)$$

In the generalized variables, from (14) the Sherwood number (Sh) is possible to be obtained:

$$Sh = \frac{kl}{D} = -\sqrt{Pe} \int_0^1 \left(\frac{\partial C}{\partial Y} \right)_{Y=0} dX, \quad (15)$$

where $C(X, Y)$ is the solution of (13) [9, 10] developed by a perturbation method [2]:

$$Sh = \sqrt{\frac{6Pe}{\pi}} \left(1 - \frac{Fo}{6} - \frac{19Fo^2}{120} \right). \quad (16)$$

The examined example shows the theoretical techniques, used for the modeling and simulation of the elementary processes in a complex industrial process. The presence of the models of all elementary processes in a complex industrial process does not give a practical opportunity to synthesize its model. The role of modeling and simulation of elementary processes is the theoretical explanation of a number of physical and chemical effects observed experimentally in industrial processes. An illustrative example in this respect is the nonlinear theory of the mass transfer.

3.2. Non-linear mass transfer theory

The theory of the diffusion boundary layer [2] is the basis of modern linear mass transfer theory, where the convection-diffusion equation in (8) is linear, i. e. the velocity does not depend on the concentration. In a number of cases, the experimental results for the mass transfer rate are higher than the predictions of the linear theory [11, 12]. This is due to nonlinear effects, where the mass transfer influences the hydrodynamics and the velocity begin to depend on concentrations. These non-linear effects are related to the induction of secondary flows at the interphase boundaries as a result of interphase mass transfer. Such effects are the effect of large concentration gradients [11], the effect of Marangoni and the effect of Stephan flow [12].

The large concentration gradients create an intensive diffusion flux that have a hydrodynamic character, and a secondary flow is induces, directed at the normal of the interphase boundary and results in an additional convective mass transfer.

The effect of Marangoni is a result of the gradient of the surface tension on the interphase surface, as a result of the surface gradient of the temperature or surface active agents concentration on the liquid-gas (liquid) interphase, and induces a tangential flow. As a result of the continuity of the flow, there appears to be a much lower flow in the direction of the normal of the interphase boundary and consequently an additional convective flow. Because of this, this effect is relatively weak and occurs in motionless or slow moving fluids.

The Stephan's flow is a result of a phase transition liquid-steam at the interphase surface when the volume of the liquid (steam) increases (decreases) a thousand times. As a result, there is a secondary flow, directed to the normal of the interphase boundary, and an additional convective mass transfer.

In the above three cases, an additional hydrodynamic effect appears very often because the secondary currents disturb the hydrodynamic stability of the flows and self-organizing dissipative structures occur, which further accelerate the mass transfer [9]. To these effects can be added the Benar instability [12] in the case of a positive vertical gradient of the density of gases or liquids resulting from concentration or temperature gradients.

The theory of mass transfer allows the construction of the process model if its physical mechanism is known. The model thus obtained allows for the identification of this mechanism, i.e. the determining of the significant physical effects and rejection of the insignificant, using the generalized analysis method [2].

4. Physical mechanism identification

The qualitative analysis of the models permits to be made the physical mechanism identification, using generalized variables [2]. They are dimensionless variables, whereas as the characteristic (inherent) scales are used the maximal or average values of the variables. The introduction of the generalized variables leads to dimensionless model. As a result the unity is the order of magnitude of all functions and their derivatives in the model, i.e. the effects of the physical and chemical phenomena (the contribution of the terms in the model), are determined by the orders of magnitude of the dimensionless parameters in the model. If all model equations are divided by the dimensionless parameter, which has the maximal order of magnitude, all terms in the model equations will be classified in three parts:

1. The parameter is unity or its order of magnitude is unity, i.e. this mathematical operator represents a main physical effect;
2. The parameter's order of magnitude is 10^{-1} , i.e. this mathematical operator represents a small physical effect;
3. The parameter's order of magnitude is $\leq 10^{-2}$, i.e. this mathematical operator represents a very small (negligible) physical effect and has to be neglected, because it is not possible to be measured experimentally.

After the physical mechanism identification the model contains a minimum number of parameters which must be determined experimentally.

5. Parameters identification

In general case, the identification of the parameters in the model is made by the minimization of the least squares function by the inverse identification problem solution [1-3]. The least squares function represents the sum of the squares of the differences between the calculated and the experimental values of the functions in the model and its minimum must be obtained with respect to the parameters in the model. This inverse identification problem is very often incorrect and needs special methods for the solutions [2].

6. Statistical analysis of the model adequacy

The stochastic nature of the errors during the experimental data determination leads to subsequent errors of the model parameters. The model is adequate if the variance of the statistical error of the model does not exceed the variance of the statistical error of the experimental data, i. e. the accuracy of the functions calculation by the model is not less than the accuracy of the function experimental measurement [2, 3].

7. Processes simulation

The process simulation uses the most common numerical methods of applied mathematics to solve model equations. For this purpose, commercial software is usually used. In many cases, however, it is necessary to introduce this software into specialized algorithms [2, 13].

8. Modeling of industrial mass transfer processes in column apparatuses

The diffusion boundary theory is not applicable for the modeling of chemical, absorption, adsorption and catalytic processes in column apparatuses, where the velocity distributions and interphase boundaries are unknown.

The use of the physical approximations of the mechanics of continua for the interphase mass transfer process modeling in industrial column apparatuses is possible if the mass appearance (disappearance) of the reagents on the interphase surfaces of the elementary physical volumes (as a result of the heterogeneous reactions) are replaced by the mass appearance (disappearance) of the reagents in the same elementary physical volumes (as a result of the equivalent homogenous reactions), i.e. the surface mass sources (sinks), caused by absorption, adsorption or catalytic reactions must be replaced with equivalent volume mass sources (sinks). The solution of this problem is related with the creation of new type of convection-diffusion and average-concentration models [13].

The convection-diffusion models permit the qualitative analysis of the processes only, because the velocity distribution in the column is unknown. On this base is possible to be obtained the role of the different physical effect in the process and to reject those processes, whose relative influence is less than 1%, i.e. to be made process mechanism identification.

The average-concentration models are obtained from the convection-diffusion models, where average velocities and concentrations are introduced. The velocity distributions are introduced by the parameters in the model, which must to be determined experimentally.

Conclusions

The theoretical foundations of modeling and simulation of the industrial processes are presented. The first step is the formulation

of the physical mechanism of the industrial process and the construction of a mathematical structure, containing the mathematical operators that quantitatively describe the individual physical effects in this mechanism. The introduction of generalized (dimensionless) variables through characteristic scales permits to be made a quality analysis of the industrial processes. The obtaining of the experimental data and using it to calculate the model parameters by solving a inverse identification problem leads to the final form of the mathematical model. The statistical analysis of the model adequacy leads to the practical applicability of the mathematical model. The presented results are published in 8 monographies (www.iche.bas.bg/Books_BG.htm).

References

1. Хр. Бояджиев, Основи на моделирането и симулирането в инженерната химия и химичната технология, ИИХ-БАН, София, 1993. (Fundamentals of modeling and simulation in chemical engineering and chemical technology - Bulgarian).
2. Chr. Boyadjiev, Theoretical Chemical Engineering. Modeling and simulation, Springer-Verlag, Berlin Heidelberg, 2010.
3. Хр. Бояджиев, Основи на моделирането и симулирането в химичната промишленост, Изд. БАН „Проф. Марин Дринов“, София, 2017. (Fundamentals of modeling and simulation in chemical industry - Bulgarian).
4. Е. Л. Фейнберг, Две культуры. Интуиция и логика в искусстве и науке, БЕК 2, Фрязино, 2004, 286 сс. (Logic and intuition in art and science - Russian).
5. Chr. Boyadjiev, Some Thoughts on Logic and Intuition in Science and Chemical Engineering, Open Access Library Journal, 1, №6, 1-5, 2014.
6. Keizer J., Statistical Thermodynamics of Nonequilibrium Processes, Springer-Verlag, New York, 1987.
7. L. D. Landau, E. M. Lifshitz, Fluid Mechanics, Pergamon, Oxford, 1989.
8. V. G. Levich, Physicochemical Hydrodynamics, Prentice Hall, New York, 1962.
9. Chr. Boyadjiev, V. Beschkov, Mass Transfer in Liquid Film Flows, Publ. House Bulg. Acad. Sci., Sofia, 1984.
10. Хр. Бояджиев, В. Бешков, Массоперенос в движущихся пленках жидкости. Москва, Изд. „Мир“, 1988. (Mass Transfer in Liquid Film Flows - Russian).
11. В. С. Крылов, Хр. Бояджиев, Нелинейный массоперенос. Новосибирск, Изд. Институт теплофизики СО РАН, 1996. (Non-Linear Mass Transfer - Russian).
12. Chr. B. Boyadjiev, V. N. Babak, Non-Linear Mass Transfer and Hydrodynamic Stability, New York, ELSEVIER, 2000.
13. Chr. Boyadjiev, M. Doichinova, B. Boyadjiev, P. Popova-Krumova, Modeling of Column Apparatus Processes, Springer-Verlag, Berlin Heidelberg, 2016.

ROTATING OF A BALL IN CHAMBER FILLED WITH A FLUID

ВРАЩЕНИЕ ШАРА В КАМЕРЕ, ЗАПОЛНЕННОЙ ЖИДКОСТЬЮ

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

Abstract: Influence of form errors of a chamber filled with a liquid on the movement and stability of a ball, rotating in the chamber ([1-3]), is studied. Two cases of the influence of a chamber form errors on the forces, acting on the ball, are defined. The first case describes the situation when limitations on the rotor shift are not imposed and disturbances of the chamber form are set by spherical harmonics not above the first order. Then the chamber of a disturbed form, from the point of view of the reaction forces of the liquid and their moments, does not differ from a similar spherical chamber. In the second case disturbance of a chamber form are arbitrary and the rotor shift is supposed small. Then the force, acting on the rotor, depends on its displacement only, and the momentum does not depend on shift. A chamber of any form is equivalent to an ellipsoid. A rising here deflective moment tends to direct the angular speed vector along the small semiaxis of the ellipsoid, i.e., a stable position of the rotor appears.

KEYWORDS: INFLUENCE OF FORM, STABILITY, FLUID, ROTATION.

1. Let us consider a Cartesian system of coordinates, the origin of which is located in the center of a rotating ball - a rotor. Otherwise, the system is arbitrary. Direction of angular speed vector $\vec{\Omega}$ is defined with the angle α (between the axis OZ and $\vec{\Omega}$ and β (between the axis OX and the projection of $\vec{\Omega}$ on the plane x, y). Also we will consider spherical coordinates (r, θ, φ) , related to Cartesian formulas:

$$x = r \sin \theta \cos \varphi, y = r \sin \theta \sin \varphi, z = r \cos \theta.$$

Let us introduce characteristic thickness of the gap $\delta = (R_2 - R_1)/R_1$, where R_2 - is the chamber radius. Then we will move on to dimensionless variables, choosing the rotor R_1 radius as a unit of distance and $1/\Omega$ is a unit of time. The equation for the rotor surface is $r = 1$ and the equation for the chamber inner surface is $r = r(\theta, \varphi)$. The problem of viscous incompressible fluid flow in the gap between the rotor and the chamber within the Stokes approximation outside the field of mass forces is the following [4]:

$$(1) \quad \Delta \vec{\omega} = 0, \quad \vec{\omega} = \text{rot} \vec{v},$$

$$\text{div} \vec{v} = 0.$$

Boundary conditions are:

$$\vec{v}|_{r=1} = -\sin \alpha \sin (\varphi - \beta) \vec{e}_\theta - [\sin \alpha \cos \theta \cos (\varphi - \beta) - \cos \alpha \sin \theta] \vec{e}_\varphi,$$

$$\vec{v}|_{r=r(\theta, \varphi)} = 0.$$

Now let us use the assumption of small thickness of the gap between the layer. To do that, we will introduce a new radial variable ξ , that is, when the inner surface of the chamber is a sphere of radius R_2 , we suppose that

$$\xi = (r - 1)/\delta.$$

Equation for the rotor surface now is $\xi = 0$, and equation for the inner surface of the chamber is:

$$\xi = h(\theta, \varphi), \quad h(\theta, \varphi) = (r(\theta, \varphi) - 1)/\delta.$$

As for the equations of motion (1) and the boundary conditions we will retain principal terms of their asymptotics only at $\delta \rightarrow 0$. Then equations (1) can be rewritten with a precision of terms of order δ .

$$(2) \quad \frac{\partial^2 \omega_r}{\partial \xi^2} = 0, \quad \frac{\partial^2 \omega_\theta}{\partial \xi^2} = 0, \quad \frac{\partial^2 \omega_\varphi}{\partial \xi^2} = 0,$$

$$(3) \quad \frac{1}{\delta} \frac{\partial v_r}{\partial \xi} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} (\sin \theta v_\theta) + \frac{1}{\sin \theta} \frac{\partial v_\varphi}{\partial \varphi} = 0,$$

$$(\Delta = \frac{1}{\delta^2} [\partial^2 / \partial \xi^2 + O(\delta)]),$$

where

$$(4) \quad \omega_r = \frac{1}{\sin \theta} \left[\frac{\partial}{\partial \theta} (\sin \theta v_\varphi) - \frac{\partial v_\theta}{\partial \varphi} \right],$$

$$(5) \quad \omega_\theta = -\frac{1}{\delta} \frac{\partial v_\varphi}{\partial \xi}, \quad \omega_\varphi = \frac{1}{\delta} \frac{\partial v_\theta}{\partial \xi}.$$

Now the boundary conditions are as follows:

$$(6) \quad v_r|_{\xi=0} = 0, \quad v_\theta|_{\xi=0} = -\sin \alpha \sin (\varphi - \beta),$$

$$v_\varphi|_{\xi=0} = -\sin \alpha \cos \theta \cos (\varphi - \beta) + \cos \alpha \sin \theta,$$

$$v_r|_{\xi=h} = v_\theta|_{\xi=h} = v_\varphi|_{\xi=h} = 0.$$

The solution of equations (2) looks like

$$\omega_\theta = A(\theta, \varphi) \xi + B(\theta, \varphi),$$

$$\omega_\varphi = C(\theta, \varphi) \xi + D(\theta, \varphi).$$

Differentiating equation (4) twice in respect to ξ and using the first equation (2) and the resultant solutions for ω_θ and ω_φ , we will obtain the following equation

$$(7) \quad \frac{\partial}{\partial \theta} (A \sin \theta) + \frac{\partial C}{\partial \varphi} = 0.$$

Now it is possible to express the components of the field of velocities v_θ and v_φ in terms of the introduced coefficients A, B, C, D , using equations (5) and the boundary equations

$$(8) \quad v_\theta = -\sin \alpha \sin (\varphi - \beta) + \delta (C \xi^2 / 2 + D \xi),$$

$$(9) \quad v_\varphi = -\sin \alpha \cos \theta \cos (\varphi - \beta) + \cos \alpha \sin \theta - \delta (A \xi^2 / 2 + B \xi).$$

Coefficients B and D can be obtained with the found above components of velocity and conditions on the chamber surface

$$(10) \quad D = \frac{1}{\delta h} \sin \alpha \sin (\varphi - \beta) - \frac{Ch}{2},$$

$$B = -\frac{1}{\delta h} [\sin \alpha \cos \theta \cos (\varphi - \beta) - \cos \alpha \sin \theta] - \frac{Ah}{2}.$$

So, angular velocity components of the fluid can be expressed in terms of coefficients A and C

$$(11) \quad v_\theta = \left(\frac{\xi}{h} - 1 \right) [\sin \alpha \sin (\varphi - \beta) + \frac{\delta}{2} Ch \xi],$$

$$(12) \quad v_\varphi = \left(\frac{\xi}{h} - 1 \right) [\sin \alpha \cos \theta \cos (\varphi - \beta) - \cos \alpha \sin \theta - \frac{\delta}{2} Ah \xi]$$

and basing on the equation of continuity and condition $v_r|_{\xi=0} = 0$

$$v_r = \frac{\delta}{\sin \theta} \left[\frac{\xi^2}{2} \left(\frac{\sin \alpha}{h^2} \frac{\partial h}{\partial \theta} \sin \theta \sin (\varphi - \beta) + \frac{\partial h}{\partial \varphi} \cos \theta \cos (\varphi - \beta) \right) - \frac{\cos \alpha}{h^2} \frac{\partial h}{\partial \varphi} \sin \theta + \right.$$

$$(13) \quad + \frac{\delta}{2} \left(\frac{\partial}{\partial \theta} (\sin \theta h C) - \frac{\partial}{\partial \varphi} (h A) \right) - \frac{\xi^3 \delta}{6} \left(\frac{\partial}{\partial \theta} (\sin \theta C) - \frac{\partial A}{\partial \varphi} \right).$$

Using condition $v_r|_{\xi=h} = 0$, we obtain one more equation for coefficients A and C

$$(14) \quad \frac{\partial}{\partial \theta} (\sin \theta h^3 C) - \frac{\partial}{\partial \varphi} (h^3 A) = - \frac{6 \sin \alpha}{\delta} \left[\frac{\partial h}{\partial \theta} \sin \theta \sin (\varphi - \beta) + \frac{\partial h}{\partial \varphi} \cos \theta \cos (\varphi - \beta) \right] + \frac{6 \cos \alpha}{\delta} \frac{\partial h}{\partial \varphi} \sin \theta.$$

We can define coefficients A and C , solving equation (14) together with equation (7), then it is possible to find the whole field of velocities.

Let us consider, first, a special case when function h does not depend on φ , that is the chamber has axial symmetry as related to axis oz . It is possible when the chamber is spherical and axis oz is directed along the line of the rotor and the chamber centres. In this case it is sufficient to differentiate the latter equation with respect to φ and substitute expression $\frac{\partial C}{\partial \varphi}$ from (7) into it. Then we will get the equation for function A :

$$\begin{aligned} \frac{\partial}{\partial \theta} [\sin \theta h^3 \frac{\partial}{\partial \theta} (\sin \theta A)] + h^3 \frac{\partial^2 A}{\partial \varphi^2} \\ = \frac{6 \sin \alpha}{\delta} \frac{dh}{d\theta} \sin \theta \cos (\varphi - \beta), \end{aligned}$$

the solution of which should be in the following form:

$$(15) \quad A(\theta, \varphi) = \frac{6 \sin \alpha}{\delta} f(\theta) \cos (\varphi - \beta).$$

As a result we obtain an ordinary differential equation for $f(\theta)$:

$$(16) \quad [\sin \theta h^3 (\sin \theta f)']' - h^3 f = h' \sin \theta$$

(here prime means a derivative with respect to θ). It is required to find solution for this equation which is continuous at $0 \leq \theta \leq \pi$. Actually, possibility to find such a solution depends on function $h = h(\theta)$, that is, on the chamber form. For a spherical chamber $h = 1 + \lambda \cos \theta$, where λ is relative displacement of centers (the distance between the centers is $|\lambda|\delta$) and the solution for equation (16) is as follows (compare [9]):

$$(17) \quad f(\theta) = \frac{\lambda}{\lambda^2 + 4} \left(\frac{1}{h} + \frac{1}{h^2} \right), \quad h = 1 + \lambda \cos \theta.$$

In the general case when h depends on φ it is possible to "integrate" equation (7) first, representing it as condition of equality of mixed second derivatives of a new function $E = E(\theta, \varphi)$:

$$(18) \quad \sin \theta A = \frac{\partial E}{\partial \varphi}, \quad C = - \frac{\partial E}{\partial \theta}.$$

The following expression can be chosen as E

$$E(\theta, \varphi) = - \int_0^\theta C(\bar{\theta}, \varphi) d\bar{\theta},$$

for which the conditions written above can be verified directly if A and C are connected with equation (7). Let us substitute this function $E(\theta, \varphi)$ into equation (14):

$$(19) \quad \begin{aligned} \frac{\partial}{\partial \theta} (\sin \theta h^3 \frac{\partial E}{\partial \theta}) + \frac{\partial}{\partial \varphi} \left(\frac{h^3}{\sin \theta} \frac{\partial E}{\partial \varphi} \right) = \\ = \frac{6 \sin \alpha}{\delta} \left[\frac{\partial h}{\partial \theta} \sin \theta \sin (\varphi - \beta) + \frac{\partial h}{\partial \varphi} \cos \theta \cos (\varphi - \beta) \right] - \frac{6 \cos \alpha}{\delta} \frac{\partial h}{\partial \varphi} \sin \theta. \end{aligned}$$

2. Let us consider the case when the chamber shape differs little from spherical. Let us set spherical form of the chamber in the form of $h_0 = 1 + \lambda \cos \theta$ (axis oz along the line of centers). Then we will set the form, which differs little from spherical, by function

$$h = h_0 + \mu h_1, \quad |\mu| \ll 1$$

and we will look for a solution of equation (19) in the form of

$$E(\theta, \varphi) = E_0(\theta, \varphi) + \mu E_1(\theta, \varphi) + O(\mu^2).$$

Substituting h and E into equation (19) and equating the coefficients at the same degrees μ . We obtain equations for definition E_0 and E_1 , where E_0 , which satisfies the condition of norming, is already known from (15), (17) and (18):

$$(20) \quad E_0(\theta, \varphi) = \frac{6 \sin \alpha}{\delta} \frac{\lambda}{\lambda^2 + 4} \left(\frac{1}{h_0} + \frac{1}{h_0^2} \right) \sin \theta \sin (\varphi - \beta).$$

Now the right part of the equation E_1

$$(21) \quad \begin{aligned} \frac{\partial}{\partial \theta} (\sin \theta h_0^3 \frac{\partial E_1}{\partial \theta}) + \frac{\partial}{\partial \varphi} \left(\frac{h_0^3}{\sin \theta} \frac{\partial E_1}{\partial \varphi} \right) = \\ = -3 \frac{\partial}{\partial \theta} (\sin \theta h_0^2 h_1 \frac{\partial E_0}{\partial \theta}) - 3 \frac{\partial}{\partial \varphi} \left(\frac{h_0^2 h_1}{\sin \theta} \frac{\partial E_0}{\partial \varphi} \right) + \\ + \frac{6 \sin \alpha}{\delta} \left[\frac{\partial h_1}{\partial \theta} \sin \theta \sin (\varphi - \beta) + \frac{\partial h_1}{\partial \varphi} \cos \theta \cos (\varphi - \beta) \right] - \frac{6 \cos \alpha}{\delta} \frac{\partial h_1}{\partial \varphi} \sin \theta \end{aligned}$$

is completely known and because it is lineary dependent on function h_1 , which can be expanded into series with respect to spherical function, it is sufficient to consider spherical functions themselves as h_1 ([5])

$$(22) \quad h_0 = P_n^m(\cos \theta) \cos m \varphi, \quad h_1 = P_n^m(\cos \theta) \sin m \varphi,$$

$$0 \leq m \leq n, \quad n = 0, 1, 2, \dots$$

Here P_n^m are adjoint Legendre functions of first kind. Let us substitute function h_1 and its derivatives into the right part of (21) and expand this part of the equation into Fourier series with respect to variable φ . If

$$E_1(\theta, \varphi) = \frac{a_0(\theta)}{2} + \sum_{m=1}^{\infty} [a_m(\theta) \cos m \varphi + b_m(\theta) \sin m \varphi],$$

then equation (21) decomposes into a finite system of ordinary differential equation. The system is finite because for functions (20), (22) in the right part of (21) a finite number of nonzero Fourier components is retained. From the condition of norming ($E(0, \varphi) = 0$) $a_m(0) = 0$. Let us look for continuous $0 \leq \theta \leq \pi$ solution for equations in the form of Fourier series with respect to the orthogonal system of adjoint Legendre functions of appropriate weight with a fixed superscript. Let us clarify what we get at small values of n , that is in the first harmonics with respect to angle φ . If $m = 0$, then $h_1 = \text{const}$ and the equation for the chamber surface takes the following form:

$$h = 1 + \lambda \cos \theta + \mu h_1 = (1 + \mu h_1) \left[1 + \frac{\lambda}{1 + \mu h_1} \cos \theta \right],$$

If $n = 1$, then for m two values are possible: $m = 0, 1$. At $m = 0$ we get

$h_1 = \cos \theta$, $h = 1 + \lambda \cos \theta + \mu \cos \theta = 1 + (\lambda + \mu) \cos \theta$ and function h is reduced to h_0 at $\lambda \rightarrow \lambda + \mu$. Let us take an arbitrary spherical function of order $n = 1$ as h_1 :

$$h = 1 + \lambda \cos \theta + \mu_1 \sin \theta \cos \varphi + \mu_2 \sin \theta \sin \varphi + \mu_3 \cos \theta =$$

$= 1 + (\lambda + \mu_3) \cos \theta + \mu_1 \sin \theta \cos \varphi + \mu_2 \sin \theta \sin \varphi$. Performing a turn, converting axis oz into axis oz' with the directing vector

$$\vec{e} = \frac{1}{\sqrt{(\lambda + \mu_3)^2 + \mu_1^2 + \mu_2^2}} \{ \mu_1, \mu_2, \lambda + \mu_3 \},$$

we will convert function h into function $h' = 1 + \lambda' \cos \theta'$, where

$$\lambda' = \sqrt{(\lambda + \mu_3)^2 + \mu_1^2 + \mu_2^2}, \quad \cos \theta' = \frac{z}{r} =$$

$$= \frac{1}{\sqrt{(\lambda + \mu_3)^2 + \mu_1^2 + \mu_2^2}} [\mu_1 \sin \theta \cos \varphi + \mu_2 \sin \theta \sin \varphi + (\lambda + \mu_3) \cos \theta].$$

So, for disturbance of the shape of chamber inner surface, which is set by spherical function h_1 of zero or first order, we have exact

solution for equation (19), deriving from (20) with simple substitution of parameters. Geometrically zero order h_1 means small extension or compression of the chamber, and the first order means a small shift with a small turn, so that the spherical shape of the chamber does not change, though the chamber undergoes some deformation and shift. Therefore for any disturbance of the chamber surface with spherical function h_1 of the order not more than one there exists an effective spherical chamber with close meanings of relative gap δ and relative shift $\lambda\delta$, giving the same values of the main forces of the fluid response to the rotor and the main vector of moment of these forces and hence leading to the same equations of the rotor motion and the same perturbing moment. Specifically, central position of the rotor equilibrium at small disturbances of the chamber shape of kind ϵh_1 remains unstable, as in the problem for the spherical chamber.

3. Let us consider a chamber of arbitrary shape which differs little from spherical. Taking into consideration that the equation for the arbitrary chamber inner surface $r = r(\theta, \varphi) > 1$ (absence of contact of rotating ball and the chamber is supposed). As a measure of the relative gap δ value, we choose average thickness of the layer with respect to the sphere

$$\delta = \frac{1}{4\pi} \int_0^{2\pi} d\varphi \int_0^\pi [r(\theta, \varphi) - 1] \sin \theta d\theta.$$

Radial variable ξ is determined as before $\xi = (r - 1)/\delta$. Then the chamber inner surface is set with equation $\xi = (r(\theta, \varphi) - 1)/\delta = h(\theta, \varphi)$. Let us suppose that function $r(\theta, \varphi)$, together with $h(\theta, \varphi)$ are twice continuously differentiated and decompose ξ into uniformly convergent series with respect to spherical functions:

$$(23) \quad h(\theta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=0}^n P_n^m(\cos \theta) (a_n^m \cos m\varphi + b_n^m \sin m\varphi),$$

where $P_n^m(\cos \theta)$ are adjoint Legendre functions [5,6]

$$P_n^m(x) = \frac{(-1)^m}{2^n n!} (1 - x^2)^{\frac{m}{2}} \frac{d^{m+n}(x^2 - 1)^n}{dx^{m+n}}, \quad 0 \leq m \leq n.$$

It is obvious from orthogonality correlation of Legendre functions and trigonometric functions the average value of function $h(\theta, \varphi)$ with respect to sphere

$$\begin{aligned} \bar{h} &= \frac{1}{4\pi} \int_0^{2\pi} d\varphi \int_0^\pi h(\theta, \varphi) P_0^0(\cos \theta) \sin \theta d\theta = \\ &= \frac{1}{4\pi} \int_0^{2\pi} d\varphi \int_0^\pi a_0^0 [P_0^0(\cos \theta)]^2 \sin \theta d\theta = a_0^0, \end{aligned}$$

but it is obvious from definition ξ and δ that $\bar{h} = 1$, that is $a_0^0 = 1$ and

$$(24) \quad h(\theta, \varphi) = 1 + \sum_{n=1}^{\infty} \sum_{m=0}^n P_n^m(\cos \theta) (a_n^m \cos m\varphi + b_n^m \sin m\varphi).$$

Now, this function should be substituted into equation (19). Let us solve it, supposing that the chamber differs little from a sphere, concentric regarding the rotor. As for geometry it means not only that the chamber shape is close to spherical, but that the rotor center is close to the chamber center, that is coefficients a_n^m and b_n^m are small. As a measure of the chamber deviation from the sphere, which is concentric regarding the rotor, we will choose function $h - 1$, equal to zero if and only if the chamber is spheric and its centre coincides with the rotor centre. We will consider as the value of this function its norm in Hilbert space L_2 of functions, which are square-integrable on the sphere

$$\|h - 1\| = \sqrt{\int_0^{2\pi} d\varphi \int_0^\pi [h(\theta, \varphi) - 1]^2 \sin \theta d\theta}.$$

Because

$$a_n^m = \frac{(2n+1)(n-m)!}{2\pi(n+m)!} \int_0^{2\pi} d\varphi \int_0^\pi [h(\theta, \varphi) - 1] P_n^m(\cos \theta) \cos m\varphi \sin \theta d\theta,$$

evaluating the latter integral with the help of Cauchy-Bunyakovskii inequality, we will obtain

$$|a_n^m| \leq \sqrt{\frac{(2n+1)(n-m)!}{2\pi(n+m)!}} \|h - 1\|,$$

that is, every coefficient a_n^m is a small value, not less than the first order regarding $h - 1$. At $a_n^m = 1$

$$\begin{aligned} &\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} (\sin \theta \frac{\partial E_n^m}{\partial \theta}) + \frac{1}{\sin^2 \theta} \frac{\partial^2 E_n^m}{\partial \varphi^2} = \\ &= \frac{3 \sin \alpha}{\delta} [(m+n)(n-m+1)P_n^{m-1}(\cos \theta) \sin((m-1)\varphi + \beta) + \\ &+ P_n^{m+1}(\cos \theta) \sin((m+1)\varphi - \beta)] \\ &+ \frac{6 \cos \alpha}{\delta} m P_n^m(\cos \theta) \sin m\varphi. \end{aligned}$$

(25)

Solution of equation (25), which is continuous on the sphere, is the following:

$$\begin{aligned} E_n^m &= -\frac{3 \sin \alpha}{\delta n(n+1)} [(n+m)(n-m+1)P_n^{m-1}(\cos \theta) \sin((m-1)\varphi + \beta) + \\ &+ P_n^{m+1}(\cos \theta) \sin((m+1)\varphi - \beta)] - \\ &- \frac{6 \cos \alpha}{\delta n(n+1)} m P_n^m(\cos \theta) \sin m\varphi. \end{aligned}$$

Solution for the equation similar to (25), but at $b_n^m = 1$, is obtained from (26) with the turn for $\pi/2$.

In terms of main force $\vec{F}$ vector definition, acting on the rotor, the chamber of arbitrary shape, which differs little from the sphere with the centre in the centre of the rotor in the first assumption regarding $h - 1$ does not differ from a spherical chamber with the same value of the gap δ , the centre of which is displaced regarding the rotor centre for the appropriate value in the appropriate direction. In particular, it is impossible to make the position of the rotor equilibrium stable, using selection of the chamber shape, if this position was unstable for the spherical chamber.

4. Let us place the beginning of Cartesian system of coordinates in the centre of spherical rotor. Let the chamber have the shape of an ellipsoid close to a sphere with semiaxis $1 + \delta_1$, $1 + \delta_2$, $1 + \delta_3$. Let us direct the axis of coordinates along the main axis of the chamber. We will define coordinates of the chamber centre as (x_0, y_0, z_0) . Then the equation for the chamber inner surface is the following:

$$(27) \quad \frac{(x - x_0)^2}{(1 + \delta_1)^2} + \frac{(y - y_0)^2}{(1 + \delta_2)^2} + \frac{(z - z_0)^2}{(1 + \delta_3)^2} = 1.$$

Let us consider $\delta_1, \delta_2, \delta_3, x_0, y_0, z_0$ to be so small that their squares and pair products can be neglected in comparison with themselves. Then equation (27) can be transformed into this equation $x^2 + y^2 + z^2 - 2(\delta_1 x^2 + \delta_2 y^2 + \delta_3 z^2) - 2(xx_0 + yy_0 + zz_0) = 1 + O(\delta_1^2 + \dots + z_0^2)$.

In spherical coordinates this equation has the following form:

$$\begin{aligned} &r^2(1 - 2\delta_1 \sin^2 \theta \cos^2 \varphi - 2\delta_2 \sin^2 \theta \sin^2 \varphi - 2\delta_3 \cos^2 \theta) - \\ &- 2r(x_0 \sin \theta \cos \varphi + y_0 \sin \theta \sin \varphi + z_0 \cos \theta) = 1 + O(\delta_1^2 + \dots + z_0^2). \end{aligned}$$

Then we will define its approximate solution, linear with respect to $\delta_1, \delta_2, \delta_3, x_0, y_0, z_0$:

$$r = 1 + \delta_1 r_1 + \delta_2 r_2 + \delta_3 r_3 + x_0 s_1 + y_0 s_2 + z_0 s_3 + O(\delta_1^2 + \dots + z_0^2).$$

Substituting r into equation (28) and equating coefficients at $\delta_1, \dots, z_0$, to zero, we will find $r_1, \dots, s_3$. So, with the accuracy up to small values of the second order:

$$\begin{aligned} r(\theta, \varphi) &= 1 + \delta_1 \sin^2 \theta \cos^2 \varphi + \delta_2 \sin^2 \theta \sin^2 \varphi + \delta_3 \cos^2 \theta + \\ &+ x_0 \sin \theta \cos \varphi + y_0 \sin \theta \sin \varphi + z_0 \cos \theta. \end{aligned}$$

Let us introduce average thickness of the gap δ according the formula: $\delta = (\delta_1 + \delta_2 + \delta_3)/3$, therefore:

$$h = \frac{r-1}{\delta} = \frac{1}{2} \left[\frac{\delta_1 + \delta_2}{2} \sin^2 \theta + \delta_3 \cos^2 \theta + x_0 \sin \theta \cos \varphi + \right.$$

$$\left. + y_0 \sin \theta \sin \varphi + z_0 \cos \theta + \frac{\delta_1 - \delta_2}{2} \sin^2 \theta \cos 2\varphi \right].$$

Decomposition of function h with respect to spherical harmonics is the following:

$$(29) \quad h = 1 + \frac{1}{\delta} [z_0 P_1^0(\cos \theta) - x_0 P_1^1(\cos \theta) \cos \varphi - y_0 P_1^1(\cos \theta) \sin \varphi + (\delta_3 - \delta) P_2^0(\cos \theta) + \frac{\delta_1 - \delta_2}{6} P_2^2(\cos \theta) \cos 2\varphi],$$

where $P_1^0(\cos \theta) = \cos \theta$, $P_1^1(\cos \theta) = -\sin \theta$, $P_2^0(\cos \theta) = (3 \cos^2 \theta - 1)/2$, $P_2^2(\cos \theta) = 3 \sin^2 \theta$; that is the chamber centre shift regarding the rotor centre, assigned with parametres x_0, y_0, z_0 , contributes to h in the form of harmonics with $n = 1$ only, though does not influence the momentum. Let us define the deviating moment $\vec{M}^p$, that is the component of $\vec{M}$, orthogonal to angle vector velocity $\vec{\Omega}$:

$$\vec{M}^p = \vec{M} - \frac{1}{\Omega^2} (\vec{M}, \vec{\Omega}) \vec{\Omega},$$

$$M_x^p = -\frac{8\pi\mu' \Omega R_1^3}{15\delta^2} [(\delta_1 - \delta_2) \sin^2 \alpha \sin^2 \beta + (\delta_1 - \delta_3) \cos^2 \alpha] \sin \alpha \cos \beta,$$

$$M_y^p = -\frac{8\pi\mu' \Omega R_1^3}{15\delta^2} [(\delta_2 - \delta_3) \cos^2 \alpha + (\delta_2 - \delta_1) \sin^2 \alpha \cos^2 \beta] \sin \alpha \sin \beta,$$

$$M_z^p = -\frac{8\pi\mu' \Omega R_1^3}{15\delta^2} [(\delta_3 - \delta_1) \sin^2 \alpha \cos^2 \beta + (\delta_3 - \delta_2) \sin^2 \alpha \sin^2 \beta] \cos \alpha.$$

It is obvious that in the ellipsoidal chamber deviation moments occur in the first approximation with regard to small parameters

$$\frac{\delta_1 - \delta}{\delta}, \frac{\delta_2 - \delta}{\delta}, \frac{\delta_3 - \delta}{\delta},$$

but for spherical chamber ($\delta_1 = \delta_2 = \delta_3 = \delta$) they are absent in the first approximation with regard to λ ($h = 1 + \lambda \cos \theta$) and occur in the second approximation with regard to λ only. It is known from the exact solution [4]. Moreover, in the case of sphere, deviating moment depends on the rotor shift. If the rotor follows the circular path, the averaged with regard to period deviating moment is equal to zero. In the case of ellipsoidal chamber deviating moment in the first approximation does not depend on the rotor position.

Let us clarify the evolution of the vector angular velocity $\vec{\Omega}$ in the simplest case of isotropic spherical rotor (density from the centre only). In this case inertia tensor is spherical and is set with a scalar I , equal to inertia moment regarding any axis, crossing the centre. Vector $\vec{\Omega}$ evolution is defined with the equation of moments $I \frac{d\vec{\Omega}}{dt} = \vec{M}$. Let us substitute components of the moment into it and obtain the system of equations

$$(30) \quad \dot{\Omega}_x = -\frac{8\pi\mu' R_1^3}{3\delta I} \left(1 + \frac{\delta_1 - \delta}{5\delta}\right) \Omega_x,$$

$$\dot{\Omega}_y = -\frac{8\pi\mu' R_1^3}{3\delta I} \left(1 + \frac{\delta_2 - \delta}{5\delta}\right) \Omega_y,$$

$$\dot{\Omega}_z = -\frac{8\pi\mu' R_1^3}{3\delta I} \left(1 + \frac{\delta_3 - \delta}{5\delta}\right) \Omega_z.$$

Its solution is the following:

$$(31) \quad \Omega_x = \Omega_x^0 \exp \left[-\frac{8\pi\mu' R_1^3}{3\delta I} \left(1 + \frac{\delta_1 - \delta}{5\delta}\right) t \right],$$

$$\Omega_y = \Omega_y^0 \exp \left[-\frac{8\pi\mu' R_1^3}{3\delta I} \left(1 + \frac{\delta_2 - \delta}{5\delta}\right) t \right],$$

$$\Omega_z = \Omega_z^0 \exp \left[-\frac{8\pi\mu' R_1^3}{3\delta I} \left(1 + \frac{\delta_3 - \delta}{5\delta}\right) t \right].$$

We can always suppose (probably, remaining the axes) that

$$(32) \quad 0 < \delta_1 \leq \delta_2 \leq \delta_3.$$

Then $\delta_3 = 3\delta - \delta_1 - \delta_2 < 3\delta$, whence it follows that:

$$0 < \frac{\delta_1}{\delta} \leq \frac{\delta_2}{\delta} \leq \frac{\delta_3}{\delta} < 3,$$

or

$$-1 < \frac{\delta_1 - \delta}{\delta} \leq \frac{\delta_2 - \delta}{\delta} \leq \frac{\delta_3 - \delta}{\delta} < 2,$$

that is addition to vector damping decrement $\vec{\Omega}$ from (31), dependent on non-spherical type of the chamber is not more than 20% in the direction of decrease and 40% in the direction of increase in comparison with a spherical chamber with the same relative thickness of the gap.

To study evolution of the vector direction $\vec{\Omega}$ we will rewrite the system of equations (30) in spherical coordinates. Therefore it is sufficient to find solution for this equation, satisfying initial condition $\beta(0) = \beta_0, 0 < \beta_0 < \frac{\pi}{2}$.

Such a solution is set by the formula

$$(33) \quad \beta(t) = \arctg[\tg \beta_0 \exp(-(A_2 - A_1)t)],$$

$$\alpha(t) = \arctg[\ctg \alpha_0 \sqrt{\frac{1 + tg^2 \beta_0}{\exp(2(A_2 - A_1)t) + tg^2 \beta_0}} \cdot \exp(-(A_3 - A_2)t)],$$

$$\alpha(t) \rightarrow \frac{\pi}{2} \text{ for } t \rightarrow \infty.$$

So, if at initial time the end of vector $\vec{\Omega}$ is in semisphere $-\frac{\pi}{2} < \beta_0 < \frac{\pi}{2}$ with the centre at the ellipsoidal chamber, then as time passes it is attracted to the end of this semiaxis, except the case of elongated ellipsoid of rotation when it is attracted to the equator along its meridian. In the opposite semisphere the same situation takes place: the end of vector $\vec{\Omega}$ evolves to the end of the small semiaxis (or, in case of elongated ellipsoid of rotation - to the equator. In this case every point of which is the end of the small semiaxis).

As a result we can conclude that in case of the ellipsoidal chamber direction of the small semiaxis is stable for vector $\vec{\Omega}$, which, under influence of deviating moment, is attracted to this direction from any initial position.

REFERENCES

1. Semenyuk, N.F., Bachinskaya N.K. "Mechanics of frictional contact of rough surfaces. The area of contact". Friction and deterioration, 1993; No. 6. 984-990.
2. Taylor, S. "Measurement and calculation of influence of a non-uniform roughness of a surface on factor of friction at turbulent current". Modern mechanical engineering, No. 7. 1989, P. 17-33.
3. Dementiev, O. "Stability of a rotor motion in a chamber filled with viscous gas". Computers & Mathematics with Application, T. 39. No. 7-8. 2000, P. 183-191.
4. Loitzanskii, L. "Mechanics of fluid and gas", Moscow, Nauka, 1978, P. 736.
5. Bateman H., Erdelyi A. "Higher transcendental functions", New York, Toronto, London, MC Graw-Hill Book comp. 1973. P.295.
6. Abramowitz, M., Stegun, I.A. Handbook of mathematical function, National Bureau of Standards Appl. Math. ser., Moscow, Nauka, 1964, P. 830.

SOFTWARE ASSURANCE OF THE SYNTHESIS AND DESIGN OF HYPERBOLOID GEAR DRIVES

Assoc. Prof. Abadjieva E. PhD.^{1,2}, Prof. Sc. D. Abadjiev V. PhD.²
Graduate School of Engineering Science, Akita University, Japan¹
Institute of Mechanics, Bulgarian Academy of Sciences, Bulgaria²

abadjieva@gipc.akita-u.ac.jp

Abstract: The study presents a brief description of the type's software, applicable to the synthesis and design of gear transmissions. The main accent is put on the approach to the computer synthesis, for which the optimization process is carried out by the method of directed search of the optimal variant of a synthesized mechanism. The specific features of the programs, applied by the authors, oriented to the synthesis of spatial gear mechanisms with linear contacting teeth with face meshing (Spiroid and Helicon) are studied in detail.

Keywords: HYPERBOLOID GEAR DRIVES, MATHEMATICAL MODELLING, SYNTHESIS, SOFTWARE ASSURANCE

1. Introduction

The contemporary requirements to the accuracy and reliability characteristics of machine products for the techniques to a great extent dictate the applied scientific methods for the technological synthesis, design and manufacture of gear drives [1 - 3]. In the processes of synthesis and design of the different types of gears, it is necessary to be solved complex set of problems, which considered all together define the desired optimal construction. In this case, an optimal construction means a gear transmission, which is capable to ensure the preliminary given kinematic and strength characteristics at the minimum cost for realization and exploitation. In essence, this is a system of requirements, related to different quality characteristics of the gear, namely [1, 4]:

- **geometric ones**, which control the kinematic exactness, smoothness of the working process, the character of the contact (placement of the contact spot; the orientation of the contact lines and the radius of the curvatures at contact points), related to the loading capacity of the gear sets and etc.;
- **dynamic ones**, which have impact on the noise and vibrations of the gear drives, the conditions for appearance of resonant phenomena and etc.;
- **strength ones** determining the durability and reliability of gear sets, including the transfer of nominal power in the process of rotations transformation with avoidance of "scoring", "pitting" and etc. on the active tooth surfaces of the synthesized mechanism;
- **economic ones**, that define the production costs (e.g., per unit of power), energy loss for the motions transformation (coefficient of efficiency), etc.

The realization of an adequate approach for the creation of the real gear drives requires this approach to be a complex one. This is consisted in considering of the required quality characteristics of the created gear mechanisms with the existing specific technological and manufacturing capabilities.

The choice of the approach should be realized in the process of synthesis and design of gear transmissions.

2. Aspects of the Computer Design of Gear Drives

The wide variety of gear drives used in industry and transport as reduction drives and multipliers, as well as the continuous pursuit of researchers to create new and improved gear mechanisms on one hand and on the other - the different and rapidly vitiating approaches to the mathematical modeling, synthesis and design make it practically impossible to create universal CAD systems. In connection with the mentioned above, a special attention should be paid to the extremely dynamic development of the modern technical computational tools and software applications. This often requires a revaluation not only of the way in which computer programs are organized, but also leads to informal changes in applied mathematical models [5]. The computer design is evolved, forming three types of software [6], in order to realize scientific studies in

the field "Theory of gearing" and to provide an adequate scientific support for this type of manufacture.

First type. The programs, included here, are designed to study the influence of the different kinematic, constructive, technological and exploitation parameters on various quality characteristics of the studied gear drives. Essentially, this type of software is not subjected to a particular strategy, associated with the design of CAD systems. The elaborated mathematical models, algorithms and computer programs are designed to determine the influence of one or other real-life existing parameters on the qualitative characteristics of the concrete gearings. However, the programs created in this case can be used as software modules, which are elements of system of criteria for quality control of the synthesized gear mechanisms.

Second type. This group is consisted of computer programs organized on the basis of algorithms, which are contained in standardization documents [7], company methodologies [8] or handbooks [9, 10]. The program products, included here, are developed on the basis of algorithms for geometric and strength calculation of the traditional types of gear drives: cylindrical involute with external and internal mating gears, cylindrical worm gears, bevel gears with straight teeth and so on. It should be noted, that the algorithms used in these cases do not ensure the optimization in the synthesis and design of gear mechanisms. Secondly, this category of software can also include those products, through which the strength characteristics of the already geometrically and technologically synthesized gear drives [7] are examined. In that capacity, these computer programs can be treated as analysis instruments of the gear mechanisms.

Third type. The computer programs included in this category are those, which are based on the mathematical models, developed on the basis of the specially oriented scientific studies. For example, for Bulgaria, these are the computer programs that deal with the synthesis and design of Spiroid and Helicon gear sets [6, 11, 12], and with conical and hypoid gear mechanism – type Gleason [13-15]; and others. For the contemporary gear transmissions, including even classical gear mechanisms, which are treated in terms of current engineering requirements, the construction of new mathematical approaches to their geometrical, technological and strength synthesis is required. The optimization synthesis process in this case is realized by application of the **method of direct search**. This method gives opportunity to reduce the number of calculated gear pairs, which compose the synthesized gear mechanism. It will be reminded, that the essence of this method is as follows:

- input parameters are defined, as well as those that will not be changed throughout the whole synthesis process;
- the variables parameters are determined as well the way of their variation, respectively;
- the process of changing of the defined variable input parameters compared to their initially given value continues, until the introduced optimization criteria are fulfilled;

- from the calculated pairs of conjugated gear sets, a final variant is chosen for which, there is the best satisfaction of the additional conditions (restrictions) introduced in the mathematical model.

In other words, the process of optimization synthesis and design of the third type of software is based on adequate iterative procedures, by which the desired solution is found by changing certain parameters.

3. Constructing of Computer Programs for Calculation of Hyperboloid Gears with a Linear Contact

The computer programs designed for the synthesis and design of linear contacting hyperboloid gear mechanisms belong to the third type software. Taking into account, the known methodological limitations when constructing this type software, the following sequence is followed, for creation of the system for the computer design of hyperboloid gears with linear contacting tooth surfaces.

3.1. Mathematical Modeling for the Synthesis and Design

When profiling of the kinematically conjugated surfaces, upon which the rotations transformation between crossed axes is carried out, the basic observed principles are the principles of T. Olivier. Thoroughly discussed in [6], it will be summarized only that part of them, which is directly related to the construction of the concrete computer programs. Two applied approaches to the construction of mathematical models for synthesis are formulated here: mathematical modeling, upon which the geometric, technologic and exploitation characteristics of the designed gear sets in a small vicinity of the pitch contact point is defined and optimized and mathematical modeling related to the ensuring of the qualitative characteristics in the entire mesh region.

It is obvious that the methodological difference between the two approaches for the synthesis of spatial gear drives requires to define in advance the adaptability of the future designed hyperboloid gear sets to one of the two approaches. The determination of the adaptability of the planned procedure for building an adequate mechano-mathematical model is a complex creative process, requiring the knowledge of both the theoretical content of the approaches to the synthesis and the specific technological and exploitation requirements characterizing the created products.

3.2. Principles of Organization of the Design Process

Here, the focused will be paid only on those principles, which are determining for the construction of computer programs for the synthesis of Spiroid and Helicon gears.

Determination of the groups of independent and variable input parameters influencing the design conditions. To the group of independent input parameters should be included a set of standardization modules, that determines the technological capabilities of the hobbing machines; coefficients that define the tooth geometry as a function of the modules; coefficients of frictions between the different pairs of materials applicable for producing of the toothing of the conjugated gear pairs; coefficients, linear and angular values associated with the design of the instrumental equipment and etc.

To the input data parameters, among which the variable ones are chosen, as a rule are included those which define the overall geometry of the calculated gear system. Here belongs the parameters determining the dimensions of the gear structure: the offset, distances from the offset to the planes in which the pitch circles lie; the angles defining the orientation of the above said planes relative to the pitch normal and etc. The variable input data include also those, from which the geometry of the conjugated active tooth surfaces depends: the independent coordinates of the tooth surfaces; their helical parameters; parameters which determine face width of the teeth, etc.

Introducing basic analytical relations, which are based on the chosen approach to the mechano-mathematical modeling. Here are included the solutions of the fundamental tasks of the synthesis upon a pitch contact point and upon a mesh region with the application of the adequate geometric interpretations of the basic equation of meshing, namely [6]: the task for the synthesis of pitch circles; the task for the definition of the geometry of the active tooth surfaces by their linear and angular characteristics in the pitch contact point; the task for defining the singularity in the pitch contact point (without describing the analytical type of the tooth surfaces); the task for analytically defining the entire mesh region; formulations of relations, which are used to determine the optimal dimensions and placement of the region of mesh on the active surface and etc. This principle of organizing the computer design includes also the introduction of geometric and kinematic relations, intended for the reduction of the input parameter sets.

Constructing the complex process for the synthesis and design of hyperboloid gear drives. This is accomplished by defining the separate stages of the synthesis and design in their sequence and interconnection. This principle, applied in the design of each computer program, is in direct dependency by the type of functioning set of criteria. Those criteria determine the defined characteristics of the quality of the gear mechanism in dependence of the accepted approach to the mathematical modelling. A distinctive characteristic of the accepted principle for construction of the complex process for synthesis and design is the chosen approach for the estimation of the calculated option of gear mechanism.

4. Software Programs for Geometric and Technological Synthesis of Spiroid Gears

The shown considerations for construction of software programs, applicable to the synthesis of spatial gear mechanisms [6, 16] are also realized when constructing three types' software products for the design of Spiroid gears, which functional relations are shown in Fig. 1. Each one of the three main directions, illustrated there, has its own importance. It means that the user can restrict himself to use the results of only one program; to analyze and interpret these results and then after an adequate assessment, to go through the entire process shown in the figure.

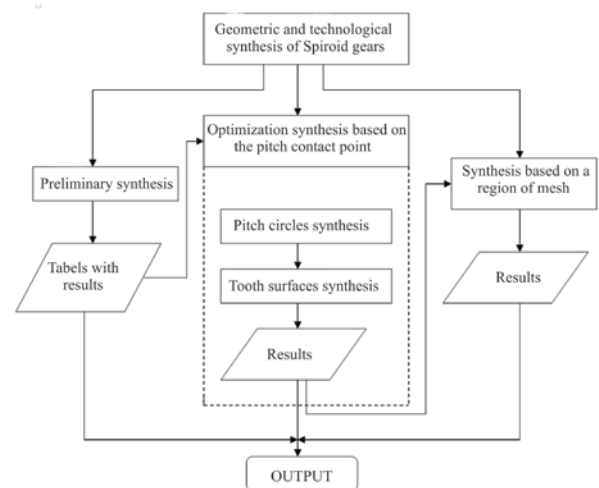


Fig. 1 General block-scheme of the approaches for the synthesis of Spiroid gears

Further bellow it will be treated only one of the directions - the central one, which includes the process of optimization geometric and technological synthesis of Spiroid gears.

4.1. Program for Optimization Geometric and Technological Synthesis of Spiroid Gears upon a Pitch Contact Point

This program consists of solving the following tasks:

- synthesis of the geometric pitch circles;

- synthesis of the active tooth surfaces of the Spiroid pinion and of the cutting tool (Spiroid hob);
- verifying the quality criteria to be fulfilled.

From the formulation of the defined tasks, it can be seen that the algorithm of this program corresponds to the approach to mathematical modeling for synthesis upon a pitch contact point. In this regard, when designing Spiroid gears, it is of particular importance to select the location of the pitch contact point in the fixed space. The placement of the pitch contact point (as a common point of the pitch circles and conjugated active tooth surfaces) affects on one hand on the common geometry of the designed gear system (overall dimensions of the gear pair) and on the other- on the geometry and proportions of the gear's teeth, as well as on the gears' quality (through the geometric, kinematic and strength characteristics of the conjugated gear pair).

For the purposes of this study, here will be briefly repeated some of the information contained in [6], which treats the geometric nature of the externally contacting geometric pitch circles. The pair of pitch circles $(H_1^c : H_2^c)$ of the mentioned type are circles H_i^c ($i = 1, 2$), having only one contact point P . Their centers are places on the axes of rotations $i-i$ ($i = 1, 2$) of the movable links, and the corresponding circles are perpendicular to those axes. Mutual position of the crossed axes is uniquely determined by the angle $\delta = \text{constant}$ between the skewed axes and by the offset (center distance) $a_w = \text{constant}$. The position of the geometric circles H_i^c ($i = 1, 2$) in respect with the rotations axes $i-i$ ($i = 1, 2$) and with their offset line is defined by four parameters for each circle. Five independent scalar equations determine the condition - that two circles have one common point. Hence, the mutual position of two circles is not defined in a unique way. It is a function of 5 independent parameters. The synthesis of geometric pitch circles is preceded by the decision, which five of the eight parameter (for both circles) to be chosen for independent ones and how to choose the intervals for their variation.

These five independent parameters, for this case, are chosen as: an angle δ between the axes of rotations $i-i$ ($i = 1, 2$); the offset a_w ; the angle δ_1 , which is concluded between H_1^c and the pitch normal $m-m$ (half of the angle at the top of the pitch cone H_1^s of the Spiroid pinion); radius r_1 of H_1^c and the distance a_1 between the offset line of the gear and the plane in which H_1^c lies.

The ambiguity of the solution enables the possibility, that these free parameters to be changed discreetly within a certain limitations and among many pairs of geometric pitch circles to look for those ones, which parameters ensure that the preliminary defined requirements for the quality of the synthesized gear drives in the vicinity of the pitch contact point are satisfied. The criteria used in the program will be discussed below.

The program allows to choose the type of conic linear helicoid, applied as active tooth surfaces of the Spiroid pinion: convolute, Archimedean or involute ones. The calculation of the necessary and sufficient geometrical and technological parameters for the design of the Spiroid gear pair and cutting instruments is realized for the required type of conic helicoid.

The indicators that serve to control the quality of the gearing are significant for the design process. As it has already be mentioned, the dependence on the solution of the task for the synthesis of the pitch circles from concrete free parameters, should be searched among the optimal geometric, kinematic and technological quality characteristics in the vicinity of the pitch contact point. They will be briefly explained.

• **The basic technological criterion.** This is the main criterion related to the technology of elaboration of Spiroid gears. It is related to the decrease in cutting tools nomenclature by ensuring the conditions for the design of Spiroid hobs with standard modules (hob parameters are functions of this module). This causes the requirement that the calculated in the pitch contact point module to coincide (with a given exactness) with any of the modules contained in the input array of standard modules.

• **Criterion that controls the singularity in the pitch contact point.** The constructed criterion is analytically described in [6, 17]. The insurance of the performance of this criterion helps to reduce the ordinary nodes from the mesh region of the synthesized gear drive. Hence, it leads to the improvement of its loading capacity, of the efficiency and of the durability. It will be reminded, that the elimination of the singularity of first order, by this criterion, is guaranteed in vicinity of the pitch contact point. The optimization, when using this criterion, is realized by the verification of the analytical dependencies introduced for each of the selected combination of the five independent variables.

• **Criterion for controlling the transmission angle of the normal force (pressure angle).** This criterion provides optimization of gear sets in terms of the transmission of normal forces from the pinion to the big gear (crown), when the gear mechanism is operated under the conditions of the rotations transformation at low-side driving.

• **Criterion for controlling the value of the Spiroid pinion spiral angle.** This criterion controls the value of this angle in the pitch contact point. Its values have to belong to definite intervals in accordance to the purpose of the design of the gear mechanism.

Here, it should specially be noted, as it is shown in [6], that the choice of the appropriate values of the pressure angle and the Spiroid pinion spiral angle of the longitudinal line of the active tooth surface $\Sigma_1^{(1)}$ in the pitch contact point substantially affect the efficiency of the gear drives. Therefore, if these geometric characteristics of the tooth surfaces of the Spiroid pinion are appropriately chosen, then an indirect control of the gear mechanism' efficiency is achieved. It also should be mentioned, that from the calculated equivalent variants of the synthesized gear mechanisms, from a geometrical and technological view point, the program allows to select that one which has the highest efficiency value for the computational (pitch) contact point.

• **Criterion related to the durability of the gear drive.** It controls the magnitude of the sliding speed at the computational contact point, depending on the chosen material for the tooth of the Spiroid gear - different types of bronze.

• **Criterion controlling the hydrodynamic conditions of meshing.** This optimization aims that the synthesized gear set has to obtain a maximum as a value -summed circumferential velocity $|\bar{V}_\Sigma|$ in the pitch contact point and minimum value of the angle Ω , which $\bar{V}_\Sigma$ concludes with the normal to the contact line in the pitch contact point.

• **Technological criteria for hobbing.** These criteria are related to the choice of the minimum value of the axial (normal) profile angle of the Spiroid hob, in order to provide optimal conditions for hobbing, both in terms of cutting the metal and in relation to the strength characteristics of the elements of the gear rack of the hob and others.

In number of cases of the design process, some of the initially independent parameters could be fixed due to the specific requirements (or example requirements for maximum sizes of the gear mechanism and the mutual position of the shafts of the gears), which results in reduction of the number of independent variables without limitation to search and find an optimal geometry of the tooth surfaces.

Input parameter of the programs are: number of Spiroid pinion threads; number of Spiroid gear teeth; offset; standard pressure angle; type of the Spiroid pinion (type of the tooth surfaces

of the Spiroid pinion); type of the bearing of the gear shafts (on two bearing supports or console); frequency of revolution and etc. Keys parameters will take values of 1 or 0 depending on whether a given criterion will be taken in consideration for the synthesis or not. For each of the free parameters should be chosen minimum and maximum values as well as the steps of variation. The independent cycles in the computer program are equal to the number of the free parameters.

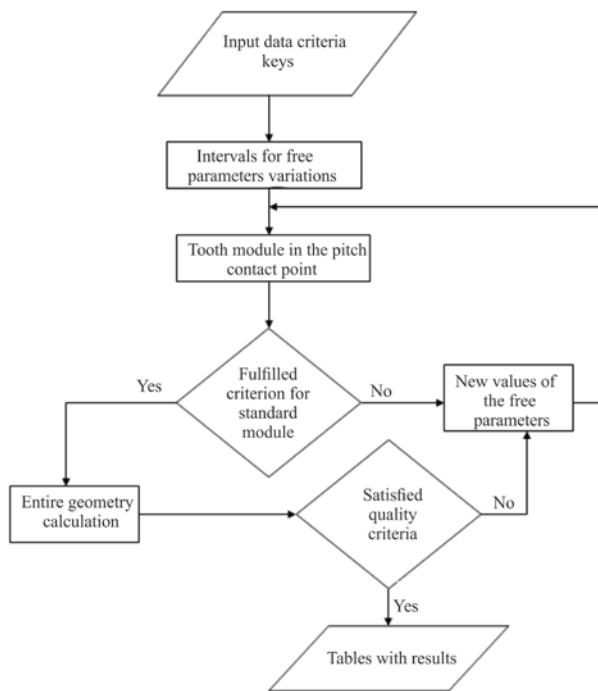


Fig. 2 Program scheme for an optimization synthesis of a Spiroid gear pair

In Fig. 2 the main block-scheme of the commented above program can be seen. The table of the results is consisted of: the basic geometric parameters of the Spiroid pinion and of the Spiroid gear, the constructive parameters of the Spiroid hob, geometric and constructive parameters of the gear pair, parameters related to the quality of meshing at the pitch contact point, such as forces in the pitch contact point, distances between bearings and pitch circles planes, efficiency, etc. It should be mentioned that, when the bearing supports are calculated, the distances between them and the pitch circles in real dimensions *mm* or dimensionless - as the ratio of the distances to the diameter of the pitch circle of the Spiroid gear should be given. In the program, all forces, that act in the pitch contact point and at the bearing supports are dimensionless in relation to the peripheral force, acting on the Spiroid gear. If in the beginning of the program, the torque on gear shaft or a torque on a pinion shaft is given, then this peripheral force has a concrete value and all forces and loads are calculated in [N]. Analogically, the sliding velocity is calculated referred to the Spiroid pinion angular velocity. If the Spiroid pinion numbers of driving motor revolutions per minute are given, then the sliding velocity is obtained in [*m/s*].

5. Conclusion

On the basis of the shown in the study basic principles of constructing computer programs for the gear drive synthesis, the accent is put on the description of the mathematical model and based on it software for an optimization synthesis upon a pitch contact point of gear pairs of type Spiroid and Helicon. The main tasks, which are solved upon this approach for synthesis and design, are shown briefly. The procedures described there are illustrated with an adequate block-scheme.

References

1. Abadjiev, V., D. Petrova. On the Synthesis and Computer Design of Spiroid Gears: A Review of the Bulgarian Approach. *Journal of Theoretical and Applied Mechanics*, Vol.33, No 3, Sofia, 2003, pp. 3-30.
2. Minkov –Petrof, K. Automated Technological Synthesis of Spatial Gearings. *Proceedings of the Sixth World Congress on "The Theory of Machines and Mechanisms"*, Volume II, December 15 – 20, 1983, Delhi, India, pp. 807-810.
3. Minkov, K., V. Abadjiev, D. Petrova. Some Aspects of the Geometric and Technological Synthesis of Hypoid and Spiroid Gears, *The Theory of Machines and Mechanisms, Proceedings of 7th World Congress of IFTOMM*, 17-22 September 1987, Seville, Spain, Pergamon Press, Volume 3, pp. 1275-1278.
4. Abadjiev, V. Bulgarian Approach to the Synthesis and, Design and Manufacture of Spiroid and Helicon Gear Drives, *Proc. from the First Conference with International Participation. Machine and Machine Elements*, 4-6 November 2004, Volume I, Sofia, 2004, pp. 70-78 (in Bulgarian).
5. Goldfarb, V. Aspects of Gears and Reduction Drives Automated Design. *Gearing and Transmissions*, Izhevsk, Assoc. Engineers, 1991, No 1, 20-24 (in Russian).
6. Abadjiev, V. Gearing Theory and Technical Applications of Hyperboloid Mechanisms, Sc. D. Thesis, Institute of Mechanics, Bulgarian Academy of Sciences, Sofia, 2007, 309p (In Bulgarian).
7. Bulgarian Standard 17108 (CT na CIV 5744 - 86). Cylindrical Involute Gears with Externally Meshing. *Standardization Strength Calculation of the Gear's Teeth*, G-15, 1999, Sofia, p. 124, (in Bulgarian).
8. Gear Engineering Standard. Bending, Stresses in Bevel Gear Teeth. Gleason Works, Rochester, N.Y., Copyright 1955, 1960, 1965, 28 p
9. Giznbrug, E., N. Golovanov, N. Firun, N. Hlebskiy. *Gears Handbook of Mechanical Engineering*. Leningrad, 1980, 416 p., (in Russian).
10. Dudley, D., J. Sprengers, D. Schroder, H. Yamashina. *Gear Motor Handbook*, Springer – Verlag, Berlin Heidelberg, 1995, 607 p
11. Abadjiev, V. On the Synthesis and Analysis of Spiroid Gears. Ph. D. Thesis, 1984, Sofia, 158 p. (in Bulgarian).
12. Abadjiev, V., D. Petrova. Aided Design of Hyperboloid Gears with High Gear Ratios. *Theoretical and Applied Mechanics*, Vol. 4, Publishing House of the Bulgarian Academy of Sciences, Sofia, 1984, pp. 24-30.
13. Anchev, A., K. Minkov, D. Petrova, Og. Kumchev. *Synthesis and Analysis of Bevel Gears with Spiral Teeth*, Publishing House of the Bulgarian Academy of Sciences, Sofia, 1980, p. 160
14. Minkov, K. Synthesis and Analysis of Semi-Conjugated Conic and Hypoid Gears. Sc. D. Thesis, LITMO, Leningrad, 1971, p. 170
15. Minkov, K. Mechanical and Mathematical Modeling of Hyperbolic Gears. Sc. D. Thesis, Sofia, 1986, 330 p. (in Bulgarian).
16. Abadjieva, E., Mathematical Models of the Kinematic Processes in Spatial Rack Mechanisms and Their Application, Ph. D Thesis, Institute of Mechanics- BAS, Sofia, Bulgaria, 2010, 165 p.
17. Abadjieva E., V. Abadjiev. Synthesis of Hyperboloid Gear Drives: Controlling of the Singularity on the Active Tooth Surfaces. *Proc. of International Conference of Science, Engineering and Technology Innovations*. March 15-17, 2017 Taipei, Taiwan, (Published on CD)

THE EFFECT OF OPTICAL AND RECOMBINATION LOSSES IN $\text{Cu}_2\text{ZnSn}(\text{S,Se})_4$ -BASED THIN-FILM SOLAR CELLS WITH CDS, ZNSE, ZNS WINDOW AND ITO, ZNO CHARGE-COLLECTING LAYERS

M.Sc. Dobrozhan O.A.*, St. Danylchenko P.S., Ph.D. Grynenko, V.V., Prof. Dr. Opanasyuk A.S.**

Department of Electronics and Computer Technology

Sumy State University, 2, Rymskogo-Korsakova st., 40007 Sumy, Ukraine

*dobrozhan.a@gmail.com, **opanasjuk_sumdu@ukr.net

Abstract: We reported the investigation the effect of the optical and recombination losses in solar cells (SCs) based on $n\text{-CdS}(\text{ZnSe}, \text{ZnS})/p\text{-Cu}_2\text{ZnSn}(\text{S,Se})_4$ heterojunctions with $n\text{-ITO}(\text{ZnO})$ frontal charge-collecting contacts on the internal (Q_{int}), external (Q_{ext}) quantum yields, short-circuit current density (J_{sc}) and maximum efficiency (η) of solar cells

KEYWORDS: $\text{Cu}_2\text{ZnSnS}_4$ FILMS; OPTICAL LOSSES; RECOMBINATION LOSSES; QUANTUM YIELD; SHORT-CIRCUIT CURRENT; EFFICIENCY.

1. Introduction

The future forecast for the renewable energy demonstrates that the solar power will become the dominant energy source from the middle of the 21st century. One of the most perspective routes to utilize the solar energy is its conversion into electricity by using the solar cells (SCs).

Nowadays, the silicon-based SCs are the most commercialized and widespread technology. But alternative are thin-film SCs based on $\text{CdS}/p\text{-CdTe}$ heterojunction (HJ) with $n\text{-ITO}$ frontal charge-collecting contact have also been widely spread on the photovoltaic market [1, 2]. However, the shortcomings, such as toxicity of Cd, high price and low abundance of In and Te, give rise to the search for the alternative materials to the functional layers in the photovoltaic devices. Nowadays, the compound $\text{Cu}_2\text{ZnSn}(\text{S,Se})_4$ (CZTSSe) is regarded as a promising substitute for the traditional Si, CdTe, $\text{Cu}(\text{In,Ga})(\text{S,Se})_2$ absorber layers. CZTSSe possesses the controlled band gap for the solar light absorption ($E_g^{\text{CZTSSe}} = 1.0\text{-}1.5$ eV), p -type conductivity and high absorption coefficients ($\alpha \sim 10^4\text{-}10^5 \text{ cm}^{-1}$) [3, 4]. The promising alternatives for the well-known SCs are considered the structure with ZnSe or ZnS window layers and ZnO charge collecting contact [5-7]. These structures contain only the abundant and non-toxic elements. ZnS, ZnO, ZnSe ($E_g^{\text{ZnS}} = 3.68$ eV, $E_g^{\text{ZnO}} = 3.37$ eV, $E_g^{\text{ZnSe}} = 2.67$ eV) are the wide-band gap semiconductors allowing to increase the number of the photons incoming to CZTSSe absorber layer.

According to Shockley-Queisser analysis, the maximum theoretical efficiency of the thin-film SCs with CZTSSe absorber layer is about (32-34) % [3, 8]. However, the experimental efficiency of CZTSSe SCs is only 12.6 % [9, 10]. The difference between the theoretical and experimental values of the efficiency can be explained by the optical, electrical and recombination losses which take place during the conversion of the solar energy into electricity.

The enhancement of the SC efficiency might be achieved by the minimization of the described losses by using the optimized structures based on the functional layers with the improved characteristics.

The foregoing discussions identified the main goal of this work – to determine and compare the optical and recombination losses in the SCs based on $n\text{-CdS}(\text{ZnSe}, \text{ZnS})/p\text{-CZTSSe}$ HJs with $n\text{-ITO}(\text{ZnO})$ frontal charge-collecting contacts.

2. Methodologies

The light flow, before reaching CZTSSe absorber layer in which the generation of the electron-hole pairs is occurred, passes through the $\text{ITO}(\text{ZnO})$ and $\text{CdS}(\text{ZnSe}, \text{ZnS})$ layers of the SCs. Herewith, the optical losses of energy as a consequence of the light

reflection from the air/ $\text{ITO}(\text{ZnO})$, $\text{ITO}(\text{ZnO})/\text{CdS}(\text{ZnSe}, \text{ZnS})$, $\text{CdS}(\text{ZnSe}, \text{ZnS})/\text{CZTSSe}$ interfaces and the light absorption of $\text{ITO}(\text{ZnO})$, $\text{CdS}(\text{ZnSe}, \text{ZnS})$ are observed.

The reflection coefficients at the interfaces of the contacted layers can be determined by using Fresnel equation [11]:

$$R = \left(\frac{n_i - n_j}{n_i + n_j} \right)^2 \quad (1)$$

where n_i , n_j – refractive indices of the first and second contacted materials, respectively.

In the case of the electrically conductive material, the reflection coefficients might be calculated by using the following expression [12]:

$$R_{ij} = \frac{|n_i^* - n_j^*|}{|n_i^* + n_j^*|} = \frac{(n_i - n_j)^2 + (k_i - k_j)^2}{(n_i + n_j)^2 + (k_i + k_j)^2} \quad (2)$$

where n_i^* , n_j^* – the complex refractive indices; k_i , k_j – the extinction coefficients.

The spectral dependencies of n and k were taken from the literature data on the refractive and extinction coefficients of ITO, ZnO, CdS, ZnSe, ZnS, CZTS [3, 13, 14]. It was assumed that the air has $n = 1$ and $k = 0$.

The transmission coefficients taking into account both the light reflection and absorption of the charge-collecting and window layers can be calculated using the expression [11, 15]:

$$T(\lambda) = (1 - R_{12})(1 - R_{23})(1 - R_{34})(1 - R_{45}) \exp(-\alpha_1 d_1) \exp(-\alpha_2 d_2) \quad (3)$$

where α_1 , α_2 – the absorption coefficients of the charge-collecting and window layers; d_1 , d_2 – the charge-collecting and window layer thicknesses.

The absorption coefficients of the materials $\alpha(\lambda)$, considering the extinction coefficient $k(\lambda)$, can be calculated by using the following equation [11]:

$$\alpha(\lambda) = \frac{4\pi}{\lambda} k \quad (4)$$

The modeling of the light reflection and absorption processes in the multilayer structures was carried out by using the different thicknesses of the window, $d_{\text{CdS}(\text{ZnSe}, \text{ZnS})} = (25\text{-}100)$ nm, and frontal charge-collecting, $d_{\text{ITO}(\text{ZnO})} = (100\text{-}200)$ nm, layers. These thickness values of the layers are typical for the practical SCs.

The important parameter for the analysis of the recombination losses in the SCs is the width of space charge region (w), in other words, the depletion region, occurring at the interface between the heteropairs, where the electrical field is acting as a separator for the photogenerated electron-hole pairs.

This width mainly depends on the concentration of uncompensated acceptors ($N_a - N_d$) (i.e., the difference between the acceptor and donor concentrations), locating in the semiconductor materials, and the contact barrier height. However, the latter value for the investigated junctions, unfortunately, was not known. This problem was solved by means of the construction of the energy band diagrams of the HJs.

It was considered that the small amount of the surface states exists at the interface between heteropairs. At the same time, the charge transport mechanism was described accordingly to Anderson model.

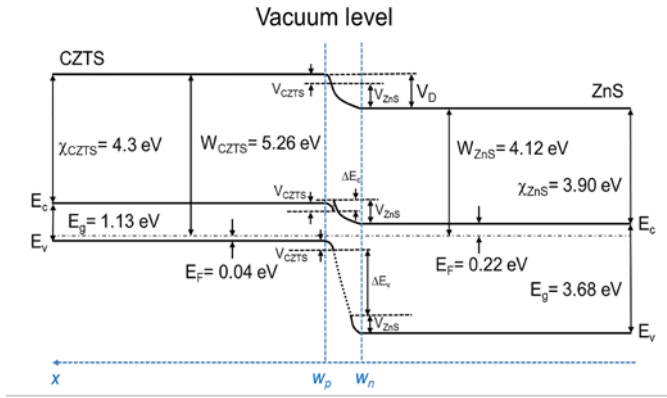


Fig. 1. Energy band diagrams of n-ZnS/p-CZTS HJ.

Unlike the fact that the charge transport processes at n-CdS/p-CdTe HJ are analogous to those occurring in the Schottky diodes [1], the same charge transport mechanisms for n-CdS(ZnSe, ZnS)/p-CZTS heterosystems are not acceptable. It is due to the fact that the doping levels in CZTS ($N_a = 10^{17}-10^{18} \text{ cm}^{-3}$ [16]) are higher than in CdTe material ($N_a = 10^{14}-10^{17} \text{ cm}^{-3}$ [17]) and even higher than in the window materials ($N_d = 10^{16}-10^{17} \text{ cm}^{-3}$ [14]). It means that SCR is located both in the window (w_n) and absorber (w_p) layers, and SCR width can be determined by the equations [18]:

$$w = \sqrt{\frac{2\varepsilon_n\varepsilon_p\varepsilon_0(V_D - qU)}{q^2} \left(\frac{1}{\varepsilon_n N_D} + \frac{1}{\varepsilon_p N_A} \right)} \quad (5)$$

where ε_n , ε_p – the relative permittivity of the window and absorber materials; ε_0 – the vacuum permittivity; $V_D = qV_{bi}$ – the contact barrier height (V_{bi} – the built-in potential); U – the applied external voltage; q – the elementary charge; N_A , N_D – the concentration of uncompensated acceptors and donors in the absorber and window layers.

Kosyachenko et al. suggest that the solution of the continuity equation is effective for the determination of the drift component of the internal quantum yield (Q_{drift}) of the SC, while taking into account the recombination at the HJ interface and in SCR, by using the following equation [11, 17]:

$$Q_{drift\ p(n)} = \frac{1 + \frac{S}{D_{p\ p(n)}} \left(\alpha_{p(n)} + \frac{2 \cdot (V_D - qU)}{w_{p(n)} \cdot kT} \right)^{-1}}{1 + \frac{S}{D_{p\ p(n)}} \left(\frac{2 \cdot (V_D - qU)}{w_{p(n)} \cdot kT} \right)^{-1}} - \frac{\exp(-\alpha_{p(n)} w_{p(n)})}{1 + \alpha_{p(n)} \cdot L_{n\ p(n)}} \quad (6)$$

where S – the recombination velocity of the charge carriers at the HJ interface and in SCR; $D_{p\ p(n)}$ – the diffusion coefficients of the holes (electrons) in the absorber (window) layers; $\alpha_{p(n)}$ – the light absorption coefficients of the absorber (window) layer; k – the Boltzmann constant; T – the temperature; $L_{n\ p(n)}$ – the diffusion

length of the electrons (holes) in the absorber (window) layer ($L_{n(p)} = (\tau_{n(p)} D_{n(p)})^{1/2}$, where $\tau_{n(p)}$ – the lifetime of the electrons (holes), $D_{n(p)}$ – the diffusion coefficients of the electrons (holes) in the relevant layers).

It should be noted that the equation (6) does not consider the recombination in the quasineutral regions of the window and absorber materials and on the back surface of CZTS layer. To account these losses, the diffusion ($Q_{diff\ p(n)}$) component of the quantum yield can be evaluated by the following equation [17]:

$$Q_{diff\ p(n)} = \frac{(\alpha_{p(n)} L_{n\ p(n)} / (D_{p\ p(n)} L_{n\ p(n)} - 1)) \exp(-\alpha_{p(n)} w_{p(n)}) + \sinh((d_{p(n)} - w_{p(n)}) / L_{n\ p(n)}) + \alpha_{p(n)} L_{n\ p(n)} \exp(-\alpha_{p(n)} (d_{p(n)} - w_{p(n)}))}{((S_n L_{n\ p(n)} / D_{p\ p(n)}) \cosh((d_{p(n)} - w_{p(n)}) / L_{n\ p(n)}) - \exp(-\alpha_{p(n)} (d_{p(n)} - w_{p(n)})) + \sinh((d_{p(n)} - w_{p(n)}) / L_{n\ p(n)}) + \alpha_{p(n)} L_{n\ p(n)} \exp(-\alpha_{p(n)} (d_{p(n)} - w_{p(n)})))}, \quad (7)$$

where $d_{p(n)}$ – the thicknesses of the absorber and window layers; S_b – the recombination velocity in the quasineutral regions and on the back surface of the absorber layer.

The total internal quantum yield of the SCs is easy to determine as the sum of all quantum yields, considering the directions of the drift and diffusion currents in the space charge and quasineutral regions. The account of the optical losses owing to the reflection and absorption of the light by the auxiliary layers (ITO, CdS, ZnSe, ZnS) of the SCs gives the opportunity to determine the external quantum yield (Q_{ext}) of the device [11, 17]:

$$Q_{ext} = T(\lambda) Q_{int} \quad (8)$$

The optical losses described in the previous subsections are important for the analysis of the SC efficiency. As a consequence of its consideration, we built the spectral dependencies of the external quantum yield (Q_{ext}) for the investigated SCs. The calculations were carried out using the following physical values: $N_a = 10^{18} \text{ cm}^{-3}$, $N_d = 10^{17} \text{ cm}^{-3}$, $d_{ITO(ZnO)} = 100 \text{ nm}$, $d_{CdS(ZnSe, ZnS)} = 25 \text{ nm}$, $d_{CZTS} = 1 \mu\text{m}$. The concentrations of uncompensated acceptors and donors coincide with SCR widths which are close to the device thicknesses. At the same time, the thicknesses of all functional layers were taken close to those used in the practical SCs [2].

The short-circuit current density (J_{sc}) of the SCs was determined using the well-known formula:

$$J_{sc} = q \sum_i T(\lambda) \frac{\Phi_i(\lambda_i)}{h\nu_i} Q_{int}(\lambda_i) \Delta\lambda_i \quad (9)$$

where $\Phi_i(\lambda_i)$ – the spectral power density of the solar radiation; $\Delta\lambda_i$ – the interval between neighboring values of the wavelength; $h\nu_i$ – the photon energy.

The calculation of J_{sc} was carried out under AM 1.5G radiation conditions [19]. Herewith, the maximum short-circuit current density ($J_{max\ sc}$) can be obtained by neglecting the light losses owing to the absorption of the auxiliary layers, i.e. $T(\lambda) = 1$, and under the circumstance that every photon generates the electron-hole pair which reaches the charge-collecting contacts without recombination, i.e. $Q_{ext}(\lambda) = 1$. It was established, that the maximum value of the short-circuit current density of the investigated SCs is equal to $J_{max\ sc} = 34.82 \text{ mA/cm}^2$.

The solar cell efficiency (η) is determined by the well-known equation [10, 29, 38]:

$$\eta = \frac{U_{oc} \cdot J_{sc} \cdot FF}{P_{in}} \quad (10)$$

where U_{oc} – the open-circuit voltage; J_{sc} – the short-circuit current density; FF – the fill factor; P_{in} – the input power (100 mW/cm^2 , illumination AM 1.5G).

To determine the effect of the optical and recombination losses on the maximum efficiencies of the SCs with ITO(ZnO)/CdS(ZnSe, ZnS)/CZTS structures, the values of open-circuit voltage were taken as those that coinciding with the height of

the contact potential differences at the HJs: $U_{oc} = (0.72 \text{ V})_{\text{CdS}}, (1.07 \text{ V})_{\text{ZnSe}}, (1.14 \text{ V})_{\text{ZnS}}$, and the values of the fill factor that matching the maximum possible $FF = 89\%$ [5]. Accordingly, it was found that the maximum efficiency of the single junction SC was 33.5% [5].

3. Results and discussions

The analysis of the optical losses owing the light reflection and absorption of the window and charge-collecting layers showed that, as it was expected, the replacement of the traditional window material (CdS) with wide band gap materials (ZnSe, ZnS) caused the increase of the transmission coefficients of the multilayer structures, primarily, in the short wave region with $d_{\text{CdS}}(\text{ZnSe}, \text{ZnS}) = (25-100) \text{ nm}$. This tendency was valid for applying ITO and ZnO layers with $d_{\text{ITO}(\text{ZnO})} = (100-200) \text{ nm}$ as the charge-collecting contacts.

ZnO layer is more attractive than ITO because it improves the light transmission coefficients toward CZTS absorber layer regardless of the considered window materials.

However, it should be noted that the values of T for the best and worst structures differed only in (5.2-13.5) %.

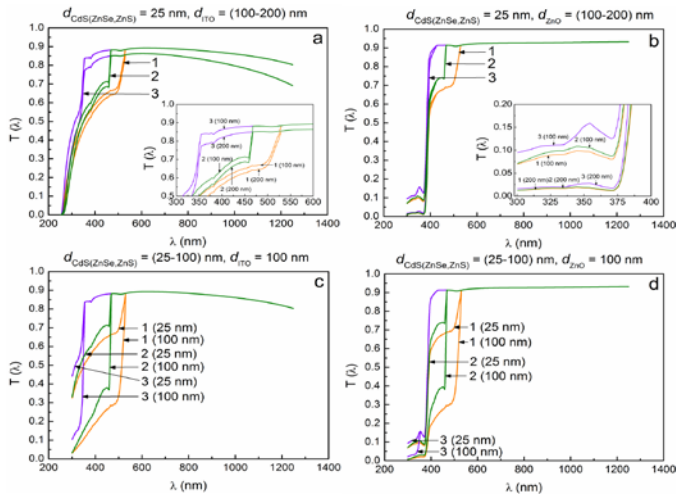


Fig.2. Spectral dependencies of the transmission coefficients of the SCs with the structures ITO/CdS/CZTS (1), ITO/ZnSe/CZTS (2), ITO/ZnS/CZTS (3) (a, c); ZnO/CdS/CZTS (1), ZnO/ZnSe/CZTS (2), ZnO/ZnS/CZTS (3) (b, d) with the different thicknesses of the window and charge-collecting layers. The light reflection from the interfaces and absorption of the auxiliary layers were taken into account.

It was established that the increase of the donor concentration in the window material at the constant values of N_a in the absorber layer resulted in the increase of the quantum efficiency of the SC based on n -CdS/ p -CZTS HJ in the photosensitive region for both CZTS and CdS materials. However, this increase had a weak influence on the internal quantum yield in the photosensitive region of the window materials in the SCs based on n -(ZnSe, ZnS)/ p -CZTS HJs. For the investigated HJs, the increase of the donor concentration caused the increase of the quantum yields in both middle and long wavelength regions, due to the extension of SCR in the absorber layer, and, as a consequence, reduced impact of the diffusion component on the total photocurrent (J_{ph}).

The analysis of the obtained dependencies showed that the values of Q_{ext} of the SC based on n -ZnS/ p -CZTS HJ were slightly higher than those of the structure with CdS and ZnSe window layers regardless of the material of the charge-collecting contacts. Thus, as it was expected, SCs with the window layers, which possess the higher values of the band gap, demonstrated the higher quantum yields.

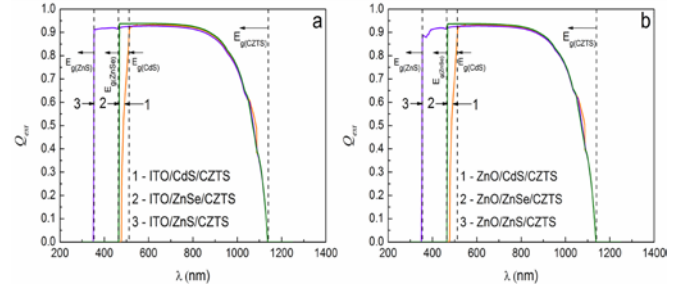


Fig. 3. Spectral dependencies of the external quantum yield (Q_{ext}) of the SCs based on n -CdS(ZnSe, ZnS)/ p -CZTS HJs with ITO (a) and ZnO (b) charge-collecting layers with: $N_a = 10^{18} \text{ cm}^{-3}$, $N_d = 10^{17} \text{ cm}^{-3}$, $d_{\text{ITO}}(\text{ZnO}) = 100 \text{ nm}$, $d_{\text{CdS}}(\text{ZnSe, ZnS}) = 25 \text{ nm}$, $d_{\text{CZTS}} = 1 \mu\text{m}$.

However, it should be noted that we neglected the inequality state at the interfaces of the different HJs. However, in reality, the mismatch density of the dislocations at the interfaces of the considered HJs is varied.

The dependencies of the SC efficiencies (η) on the thicknesses of the window (CdS, ZnSe, ZnS) and charge-collecting (ITO, ZnO) layers are presented in Fig. 3. As can be seen from Fig. 4, the best devices,

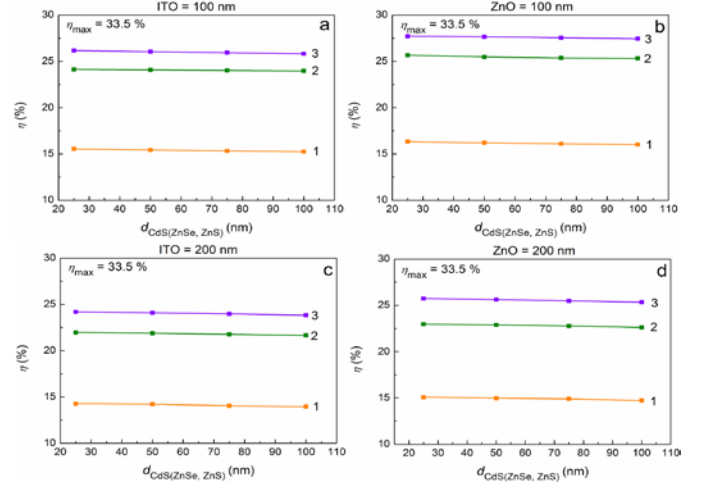


Fig.4. The effect of the optical and recombination losses on the efficiency of the SCs based on CdS/CZTS (1), ZnSe/CZTS (2), ZnS/CZTS (3) HJs with the variable thicknesses of the window layers and two constant thicknesses of the charge-collecting contacts: 100 nm (a, b) and 200 nm (c, d).

among the investigated SC structures, contain ZnS window layer ($\eta = 23.8-27.7\%$), and the highest values of the efficiency were demonstrated by a device with ZnO/ZnS/CZTS structure ($\eta \sim 28\%$ with $d_{\text{ZnO}} = 100 \text{ nm}$, $d_{\text{ZnS}} = 25 \text{ nm}$).

It should be mentioned, that the efficiency of the well-known SC with ITO/CdS/CZTS structure was about (13.9-15.5) %. These values are well correlated with the results obtained for the best SC with the analogous structure ($\eta = 12.6\%$) [9]. The SCs with ZnSe window layer showed the quite high efficiencies as well, $\eta = (21.7-25.7)\%$.

4. Conclusions

It was found that, under the consideration of the losses owing to the reflection and absorption of the auxiliary layers of devices, the values of J_{sc} of the SCs with the ZnO/ZnS/CZTS ($d_{\text{ZnS}} = (25-100) \text{ nm}$, $d_{\text{ITO}(\text{ZnO})} = 100 \text{ nm}$) structure were higher in (3.06-3.27) mA/cm^2 than those obtained for devices with ITO/CdS/CZTS structure in the overall interval of the thickness alteration. The increase of the charge-collecting layer thickness up to 200 nm led to the decrease of J_{sc} and the difference between the best (ZnO/ZnS/CZTS) and worst (ITO/ZnSe/CZTS) structures of the SCs was found to be $\sim 3.15 \text{ mA/cm}^2$. It should be noted that the

optical and recombination losses caused the decrease of J_{sc} by (21.5-37.4) %.

The best devices, among the investigated SC structures, contain ZnS window layer ($\eta = 23.8$ -27.7 %), and the highest values of the efficiency were demonstrated by the device with ZnO/ZnS/CZTS structure ($\eta \sim 28$ % with $d_{ZnO} = 100$ nm, $d_{ZnS} = 25$ nm). The SCs with ZnSe window layer showed the quite high efficiency as well, $\eta = (21.7$ -25.7) %. It should be mentioned, that the efficiency of the well-known SC with ITO/CdS/CZTS structure was about (13.9-15.5) %. These values are well correlated with the results obtained for the best SC with the analogous structure ($\eta = 12.6$ %).

The presented results show the maximum values of the efficiencies of the SCs based on n -CdS(ZnSe, ZnSe)/ p -CZTS HJs and open the way for the optimization of the practical thin film SCs.

5. References

- [1] S. M. Sze, and K. K. Ng, *Physics of Semiconductor Devices*, John Wiley & Sons, Hoboken (2006).
- [2] S. Girish Kumar, and K. S. R. Koteswara Rao. *Physics and Chemistry of CdTe/CdS Thin Film Heterojunction Photovoltaic Devices: Fundamental and Critical Aspects*, Energy Environ. Sci. vol. 7, 45 (2014).
- [3] K. Ito, *Copper Zinc Tin Sulfide-Based Thin Film Solar Cells*, John Wiley & Sons, Chichester (2015).
- [4] M. P. Suryawanshi, G. L. Agawane, S. M. Bhosale, S. W. Shin, P. S. Patil, J. H. Kim, and A. V. Moholkar. CZTS Based Thin Film Solar Cells: A Status Review. *Mater. Technol.* vol. 28, 98 (2013).
- [5] M. Nguyen, K. Ernits, K. F. Tai, C. F. Ng, S. S. Pramana, W. A. Sasangka, S. K. Batabyal, T. Holopainen, D. Meissner, and A. Neisser. ZnS Buffer Layer for $Cu_2ZnSn(SSe)_4$ Monograin Layer Solar Cell. *Sol. Energy*. vol. 111, 344 (2015).
- [6] H. Katagiri, K. Saitoh, T. Washio, H. Shinohara, T. Kurumadani, and S. Miyajima. Development of Thin Film Solar Cell Based on Cu_2ZnSnS_4 Thin Films. *Sol. Energy Mater. Sol. Cells*. vol. 65, 141 (2001).
- [7] O. Dobrozhan, A. Opanasyuk, M. Kolesnyk, M. Demydenko, and H. Cheong. Substructural Investigations, Raman, and FTIR Spectroscopies of Nanocrystalline ZnO Films Deposited by Pulsed Spray Pyrolysis. *Phys. Status Solidi A*. vol. 212, 2915 (2015).
- [8] A. Polman, M. Knight, E. C. Garnett, B. Ehrler, and W. C. Sinke. Photovoltaic Materials: Present Efficiencies and Future Challenges. *Science*. vol. 352, aad4424 (2016).
- [9] W. Wang, M. T. Winkler, O. Gunawan, T. Gokmen, T. K. Todorov, Y. Zhu, and D. B. Mitzi. Device Characteristics of CZTSSe Thin-Film Solar Cells with 12.6% Efficiency. *Adv. Energy Mater.* vol. 4, 1301465 (2014).
- [10] M. A. Green, K. Emery, Y. Hishikawa, W. Warta, and E. D. Dunlop. Solar Cell Efficiency Tables (Version 48). *Prog. Photovoltaics*. vol. 24, 905 (2016).
- [11] L. A. Kosyachenko, E. V. Grushko, and X. Mathew. Quantitative Assessment of Optical Losses in Thin-Film CdS/CdTe Solar Cells. *Sol. Energy Mater. Sol. Cells*. vol. 96, 231 (2012).
- [12] L. A. Kosyachenko, X. Mathew, P. D. Paulson, V. Ya. Lytvynenko, and O. L. Maslyanchuk. Optical and Recombination Losses in Thin-Film $Cu(In,Ga)Se_2$ Solar Cells. *Sol. Energy Mater. Sol. Cells*. vol. 130, 291 (2014).
- [13] S. O. Kasap, *Optoelectronics and Photonics: Principles and Practices*, Prentice Hall, Upper Saddle River (2001).
- [14] S. Adachi, Editor, *Handbook on Physical Properties of Semiconductors*, Kluwer Academic Publishers, Berlin (2004).
- [15] T. Mykytyuk, V. Y. Roshko, L. Kosyachenko, and E. Grushko. Limitations on Thickness of Absorber Layer in CdS/CdTe Solar Cells. *Acta Phys. Pol. A*. vol. 122, 1073 (2012).
- [16] J. P. Teixeira, R. A. Sousa, M. G. Sousa, A. F. da Cunha, P. A. Fernandes, P. M. P. Salomé, and J. P. Leitão. Radiative Transitions in Highly Doped and Compensated Chalcopyrites and Kesterites: The Case of Cu_2ZnSnS_4 . *Phys. Rev. B*. vol. 90, 235202 (2014).
- [17] L. A. Kosyachenko. Problems of Efficiency of Photoelectric Conversion in Thin-Film CdS/CdTe Solar Cells. *Semiconductors*. vol. 40, 710 (2006).
- [18] M. Balkanski, and R. F. Wallis, *Semiconductor Physics and Applications*, Oxford University Press, Oxford (2000).
- [19] Reference Solar Spectral Irradiance at the Ground at Different Receiving Conditions, Part 1: Direct Normal and Hemispherical Solar Irradiance for Air Mass 1.5. Standard of International Organization for Standardization, ISO 9845-1:1992.

INTEGRAL ASSESSMENT OF ENVIRONMENTAL QUALITY AND THE QUALITY OF LIFE OF THE POPULATION OF THE ARCTIC REGIONS OF RUSSIA IN THE PERIOD FROM 2003 TO 2015

Prof. Dr. V.V. Dmitriev

Institute of Earth Sciences – Saint Petersburg State University, Saint Petersburg, Russia

E-mail: vasily-dmitriev@rambler.ru

Abstract: *The integral assessment of environmental quality and quality of life of the population of 9 regions of the Arctic zone of the Russian Federation in the period from 2003 to 2015 is considered. To build integrated indicators, we used: summary indicators, randomized summary indicators, and "ASPID methodology" (analysis and synthesis of indicators in the information deficit). When calculating weights, incomplete, inaccurate, non-numeric information was taken into account) about the criteria and priorities of the evaluation. To assess the quality of the environment, 8 parameters were used. To assess the quality of life of the population, the state of three subsystems was taken into account: ecological (8 parameters), economic (5 parameters), social (5 parameters). The choice of criteria was made taking into account the information available on the website of the Federal Service of State Statistics of the Russian Federation, in the collections "Regions of Russia" and in state reports "On the state of the environment ...". To assess the quality of the environment in the regions and the quality of life of the population of the regions, five quality classes were introduced (I - high, II - above average, III - average, IV - below average, V - low). In constructing integral indicators, the sum of the normalized values of the indicators within subsystems (blocks) and between them was used as a synthesizing function, taking into account the equilibrium or nonequilibrium setting of priorities. When assessing the quality of the environment, all regions fall into the third class (middle - the right border of the class) with a slight temporal change. In evaluating the quality of life, three groups of regions were identified. In the first group, the quality of life for the period under review improved by 10 percent or more. The second group includes regions with an improvement in the quality of life by 5-10%, the third group includes regions with an improvement in the quality of life up to 5%. For the same time interval, the quality of life of the APR regions was compared with the regions of Central Russia (Tver Region). The forecast scenarios of a possible change in the quality of the environment and the quality of life of the population in the regions are considered. The studies were carried out with the support of the RFBR grant No. 16-05-00715-a.*

Keywords: INTEGRATED ASSESSMENT, QUALITY OF ENVIRONMENT, QUALITY OF LIFE, ARCTIC REGIONS

1. Introduction

The relevancy of the work is accounted for by the necessity to develop the theory and practice of the evaluation of the state of complex systems in nature and society, their non-additive (emergent) properties, and the system simulation of natural and social transformation of eco-, geo-, and socio-systems. The recent recommendations in the sphere of global, regional, economic, and social development are specified in the works of the commission of Stiglitz – Sen – Fitoussi (see www.stiglitz-sen-fitoussi.fr) [1]. The Commission is more known as the Stiglitz Commission, was set up in February 2008 at the initiative of the President of the French Republic under the guidance of the Nobel prize winner in economics Joseph Stiglitz, with participation of the Nobel prize winner Amartya Sen. The commission was authorized to identify the suitability of using the existing national indicators of development and progress, including such indicator as gross domestic product (GDP). It was necessary to validate the economic development and social progress parameters, to study what additional information could be required to form a more adequate picture, to discuss how to present this information correctly, and to check the feasibility of the suggested tools of measurement. The Commission submitted its first report on 14 September 2009 and called upon national statistics authorities under the aegis of international ones to focus their efforts on the development of new indicators of social progress for more adequate assessment of quality of people's life in countries and regions. To arrange its activities, the Commission was broken into three working groups which studied respectively the traditional GDP evaluation issues, life quality and sustainability issues. The working groups submitted recommendations for each of these spheres [1], which have come to be known as 12 recommendations for the fundamental amendment of the state statistics basics in France and the entire world.

The main conclusions of the report E/CN.3/2011/1 of the UN National Institute of Statistics and Economic Studies in 2011 (section B. Basic conclusions of the report) state that "well-being includes both economic resources such as income, and non-economic aspects of people's life (what they do and what they can do, how they feel, in what natural environment they live)". The

sustainability of these levels of well-being depends on our ability to pass on to future generations the accumulated assets which are significant for our life (natural, physical, human, social). Therefore it is important to discriminate between evaluation of the current well-being and evaluation of its sustainability in time" [1, p.2].

It is in this connection that the resolutions can be mentioned which were passed by the UN General Assembly on 25 September 2015 "Transforming our world: Agenda of sustainable development for the period until 2030" and "Report of the interdepartmental group of experts for indicators of objective fulfillment in the sphere of sustainable development" at the UN forty-seventh session 8-11 March 2016 [2,3].

The specific feature of the modern stage is not only the validation of the representative criteria or groups of criteria for the evaluation of the state of natural and socio-ecological-economic systems, but also the development of models of analysis and synthesis of indicators taking into account the use of incomplete, inaccurate, non-numerical information on the evaluation criteria and priorities [4].

The article discusses evaluation of the state of socio-ecological-economic systems (SEES) and quality of life of people of the RF regions. The state of SEES is believed to be the characteristic of the system at a certain moment of time. The focus is on the comparative evaluation of the quality of life in the regions of the Arctic zone of the RF (AZR) from 2003 to 2015.

The key point of our publications on the integral evaluation of the state of eco-, geo-, socio-systems and their emergent properties [6-10] is the following conclusion: in the multi-criteria evaluation of the state of the systems with an indicator approach the incomparability of the obtained assessments is revealed when according to one criterion (indicator) or a group of criteria the system is referred to one class, and according to another (others) it is referred to another (other) class (-es). Thus, with the indicator evaluations of SEES states uncertainties arise in the treatment of the obtained results. The authors have to write on what number of criteria a system can be referred to each of the classes, and more frequently, without introducing classes, the results are just ranked for each of the indicators, determining the place of SEES in question in the list of similar national or regional systems. In the

same way the SEES state is evaluated by each of the indicators recommended by them or national statistics authorities. The indicators are not generalized, and if they are, it is on the additive (grade) basis, without taking into account the priority of indicators or their trustworthiness. Then, it is often noted that the objective information of national statistics authorities is unsuitable for evaluation, because it does not contain a number of indicators already tested abroad or recommended recently. At that, the indicators taking into account the perception by people of their position in society, i.e. subjective indicators, are often used as the basic ones. We noted that in order to take such indicators into account, it is logical to use non-numerical (ordinal), inaccurate (interval), incomplete information (so-called *nnn-information*) which is necessary to take into account both for specifying the indicators and for determining the evaluation priorities. It was recommended to use multi-criteria and multi-level evaluation accounting for simulation of evaluation priorities inside levels (subsystems) and between them on the basis of “nnn-information”. The levels can include groups of criteria based on the data of national statistics authorities, as well as subjective data obtained in the statistical polls of the people. The following is recommended as the methods: a method of consolidated indicators (MCI), a method of randomized consolidated indicators (MRCI) and its modern version named by the author “ASPID methodology”, the methodology of analysis and synthesis of indicators in the information deficit [4]. In all cases a possible change in the priorities for evaluation inside the groups and between them is taken into account.

2. Experimental procedure

The modern foreign level of research is characterized by the developed methods of analysis of target indicators used to characterize the state of complex systems (mostly economic or socio-systems and their subsystems) and, to a less degree, their emergent properties. The present time is characterized by accumulation of methodological and practical experience in the research of the state of complex systems in nature and society and their separate subsystems. The method of making up a “web diagram” (“rose diagram” in Russian publications) is often used which units in a single picture the information on a great number of indicators [5]. As a result, there is an analog of the natural Hutchinson niche or socio-ecological-economic niche visually characterizing the aggregate of the conditions of existence of the system. In case of transformation system its visual image, a niche, is also transformed and its GIS-image reflects the result of impact on the system (its area (volume) is changed, the system acquires a new, predominant vector of development which is revealed and visualized graphically).

The other approach [5] uses the method of building a composite indicator that is a union of the aggregate of the used parameters into a single composite indicator (composite sustainability indicator - CSI) in which the parameters are taken into account with their weights reflecting the priority of each of them (Fig.1).

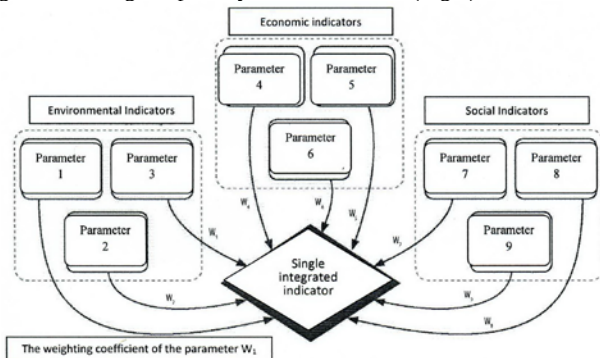


Fig.1 The structure of the integrated stability indicator (by Noam Lior, 2015)

Thus, the building of an integral indicator is implemented at one level of convolution of indicators as a sum or product of the characteristics taken with their weights. At that, the author does not consider peripheral issues of the rating of indicators taking into account the type of connection and its non-linearity; simulation of priorities (weighting factors) of the characteristics; creation of the effect of hierarchy on account of introduction of multi-level convolutions; investigation with the use of integral indicators of emergent properties of the systems (stability, well-being etc.).

In our case the building of classification models containing several levels of investigation and convolutions of indicators (Fig.2) is implemented. This figure represents one of the models that we used to evaluate the quality of life in the RF regions [6,7]. These works and Fig.2 give the units, the parameters of each unit and the results of evaluation of the quality of the environment and the quality of life of people of the arctic regions of the Russian Federation. All indicators were sampled from the data of the Rosstat website (“Regions of Russia” collections) for the period from 2003 to 2015. The basic objective of the investigations was to perform a convolution of indicators at the first and second levels and the identification of the situations in which SEES cannot retain its properties and mode parameters with a certain hypothetic influence on it in separate subsystems and the system as a whole [6].

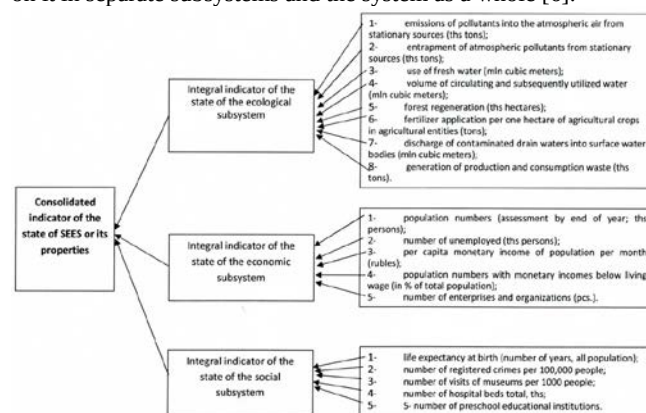


Fig. 2 Building a model for integral assessment of the SEES state, quality of the environment, and quality of life of the region's population[6].

The integral indicator Q_i was constructed so that it depended not only on the rated values of the initial characteristics q_i , but also on their priorities determined by the weights p_i , the sum of which should equal 1.0 ($0 \leq p_i \leq 1$). As an expression for the integral indicator, the linear convolution was used in the form of: Q_i

$$= \sum_{i=1}^n q_i p_i, n \text{ is the number of the evaluation criteria. The state of}$$

the system and the quality of life of people of the region were evaluated for 5 classes (I – high; II – above average; III – average; IV – below average; V – low). The proximity of the integral indicator to 0.0 evidenced high quality of life of people, the proximity to 1.0 evidenced low quality.

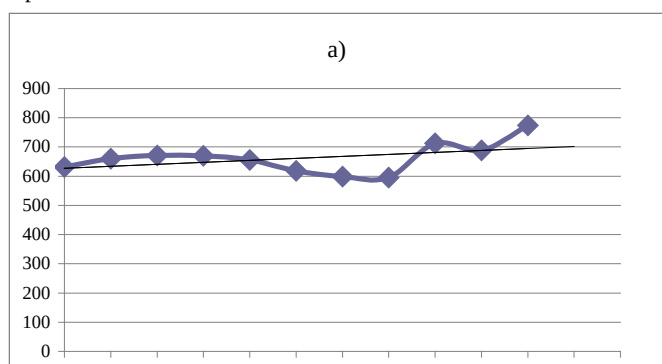
In [6,7] we investigated a change in the quality of life in 9 regions of the arctic zone of the Russian Federation (AZR) for 10-12 years, and the results of the experiments in a hypothetic change of the situations in each of the units and in all units simultaneously were described. As a result, the integral indicators of the quality of life of people for 8 scenarios for the first (inside the subsystems) and the second (between the subsystems) levels of the convolution of the indicators was calculated. The results of the evaluation of the quality of the environment and the quality of life were compared for 2003 and 2013. In these models the liner rating functions were used in the rating of the indicators with the equal weight of the evaluation parameters inside the three subsystems (ecological, economic, social) and between them. In the experiments with loads the results of the options with 30% and two-time deterioration of

the situation inside and between the units against the background of 2013 were described.

In general, the subsystem of social conditions was found to be the most sensitive subsystem. The maximum increase in the effect of impact of both on separate subsystems and on the socio-ecological-economic system as a whole (summary evaluation). With

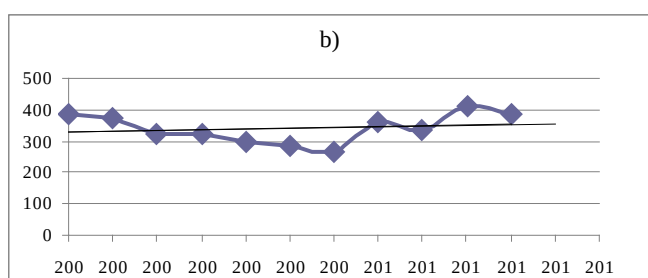
3. Results and discussion

In [7] the main drawback of the experiments with a hypothetical change in the situations in the regions was noted. It consisted in the fact that under real conditions it is not logical to expect a simultaneous change in the load by 30%, 50%, two times, etc. inside one of the units or in all the subsystems simultaneously. Each parameter chosen as a representative criterion will have its rates and direction of the changes. The situation is complicated by the different rates and direction of such changes that may be noted in different regions. Therefore, in the next stage of the investigations it was necessary to study the temporal change of each of 18 criteria and obtain the trends of these changes for the regions. Fig.3 (a,b,c,d) shows the examples of such changes for four parameters of the environmental quality subsystem (ecological unit) for the Republic of Komi for 2003-2013.

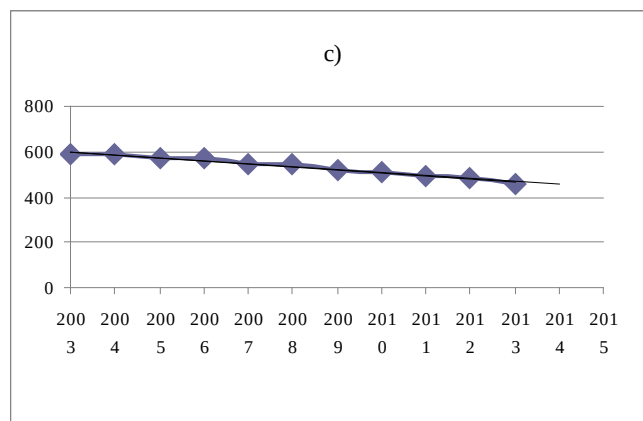


a) Contaminants emitted into atmosphere from stationary sources (thousands tons/year)

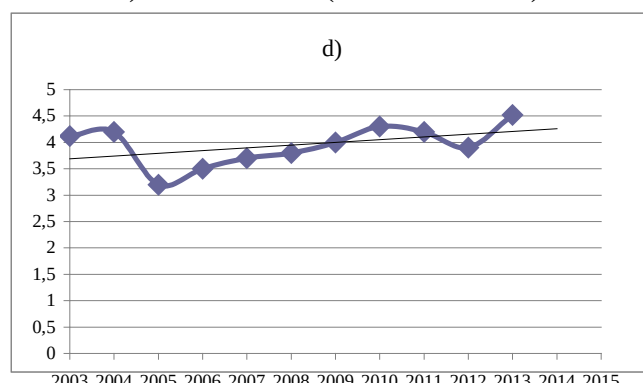
Fig.3 Tendency of changes in certain parameters of the evaluation of the quality of the environment for the Republic of Komi for 2003-2013.



b) Entrapment of atmospheric contaminants emitted by stationary sources (thousands tons/year)



c) Use of fresh water (million cubic meters)



d) Fertilizer application per one hectare of sowed crops in agricultural crops in agricultural entities (tons)

Analysis of the trends of change of separate characteristics, as was expected, showed that the rates and the direction of their change are different. Thus, for the ecological unit for the region of the Republic of Komi for 5 parameters out of 8 in 2014 a decrease in the indicators in percent compared to 2013 (indicator No. is given in Fig.2) was expected by: 1-9,5; 2-8,3; 4-0,5; 5-5,5; 6-6,0. For the rest of the parameters an increase in the indicators compared to 2013 was obtained by: 3-3,1; 7-1,9; 8-54,9. Thus, the maximum increase in respect of 2013 is noted for parameter No.8, generation of production and consumption waste, by 54.9%, the minimum decrease is noted for parameter No.4, the volume of the recycled and successively utilized water, by 0.5%.

The forecast change of the integral indicator of the ecological unit for 2014 with the quality of the weights yielded the value of the integral indicator $Q_I=0,49$, which allowed the quality of the environment to be referred to quality class III with the width of the class interval 0.37-0.56. The statistical reporting data (Rosstat website, "Regions of Russia" collections) for 2014 and 2015 confirmed the forecast calculations of the integral indicators for these years. The integral indicators of the quality of life of people in the AZR regions for the period from 2003 to 2015 (the second level of the convolution of the indicators) are given in Table 1. The scale of an integral indicator of the second level of convolution for equal priorities at the first and second levels: I – high (0-0.16); II – above average (0.16-0.36); III – average (0.36-0.56); IV – below average (0.56-0.79); V – low (0.79-1).

As a result, we note that in 8 regions there is a tendency for improvement of the quality of life of people. In Murmask region, the Republic of Komi, Khanty-Mansiysk Autonomous Okrug Yugra, Republic of Sakha (Yakutia) there is an improvement of the

quality of life of people by 7-10%. In Arkhangelsk region, Nenets Autonomous Okrug, Chukotka Autonomous Okrug, Yamalo-Nenets Autonomous Okrug the improvement of the quality of life of people is by 10-12%. The unchanged is the quality of life in Taymyr Dolgano Nenets Autonomous Okrug in the period from 2003 to 2005 (Table 1).

To compare the quality of life of the Arctic regions of the RF with the regions of the central part of the Russian Federation, the

quality of life of people in Tver region for 2003 and 2013 was examined. By the value of the integral indicator the quality of life of people in Tver region from 2003 (0.64) to 2013 (0.57) was improved by 11%. This indicates the close rates of changes in the quality of life in the regions being compared.

Table 1: Integral indicators of quality of life of people in the regions of the arctic zone of the RF for the period from 2003 to 2015 (the second level of the convolution of indicators).

Region / Year	2003	2005	2010	2013	2015
Arkhangelsk region	0,65 (IV)	0,64 (IV)	0,61 (IV)	0,58 (III - IV)	0,57 (III - IV)
Murmansk region	0,65 (IV)	0,65 (IV)	0,61 (IV)	0,58 (III - IV)	0,60 (IV)
Nenets Autonomous Okrug	0,66 (IV)	0,62 (IV)	0,59 (IV)	0,63 (IV)	0,55 (III)
Taymyr Dolgano Nenets Autonomous Okrug	0,66 (IV)	0,66 (IV)	-	-	-
Chukotka Autonomous Okrug	0,63 (IV)	0,61 (IV)	0,60 (IV)	0,57 (III - IV)	0,55 (III)
Republic of Sakha (Yakutia)	0,62 (IV)	0,61 (IV)	0,59 (IV)	0,55 (III)	0,56 (III - IV)
Yamalo-Nenets Autonomous Okrug	0,63 (IV)	0,65 (IV)	0,60 (IV)	0,57 (III - IV)	0,55 (III)
Republic of Komi	0,67 (IV)	0,67 (IV)	0,62 (IV)	0,60 (IV)	0,61 (IV)
Khanty-Mansiysk Autonomous Okrug Yugra	0,60 (IV)	0,64 (IV)	0,60 (IV)	0,54 (III)	0,55 (III)

Note. 1. The table gives the value of integral indicator; in brackets the class of quality of life; 2 – for Taymyr (Dolgano-Nenets Autonomous Okrug) the data prior to 2005 are given, for by the results of the referendum held on 17 April 2005, from 1 January 2007 Taymyr (Dolgano-Nenets) Autonomous Okrug was abolished, and the municipal Taymyr Dolgano-Nenets Autonomous region was made part of Krasnoyarsk Territory as an administrative-territorial unit with a special status.

4. Conclusions

To conclude, we will note the advantages of using the examined approach for evaluation of integrative properties of complex natural and social systems and the quality of life of people. When building classification models, the investigator introduces the classes of states of the systems and the quality of life; uses the axiological approach and axiometry (ecological qualimetry), validates the type of an integral indicator, solves the problem of rating initial data taking into account the type of communication (direct, reverse) and its non-linearity, takes into account *nnn*-information on evaluation priorities; works with evaluation scales of the necessary and sufficient evaluation criteria, though may also use qualitative scales for evaluation; can introduce several levels of convolution of indicators, specifies or simulates weights (priorities) of evaluation inside groups, subsystems (levels) between them and can change them when necessary.

The use of models in GIS environment allows territories to be divided or zoned by values of integral indicators; the temporal dynamics and spatial differences of integral indicators to be traced, a conclusion on the ability of the systems to safeguard or change the class of state (quality) in time to be drawn. The flexibility of the model algorithms allows a hierarchical system of multi-level evaluation of the state of systems when there is an uncertainty at each level of hierarchy to be organized, and the mapping of integral indicators, division and zoning of territories on this base determines new capacities of the evaluation of states of regions and quality of life of their population.

5. Acknowledgements

This research was financially supported by the Russian Foundation for Basic Research by the following grant projects: 16-

05-00715-a“Development and testing of models of integrated assessment of the sustainability of land and aquatic landscapes and socio-ecological-economic systems”.

References

- [1] United Nations Organization. Economic and Social Council. Statistical Commission Forty-second session 22–25 February 2011 Clause 4(o) of the preliminary agenda E/CN.3/2011/1. Questions for information: measuring efficiency of economics and social progress. Report of National Institute of Statistics and Economic Studies. I. Recommendations of Stiglitz Commission. P.2.
- [2] Transforming our world: Agenda of sustainable development for the period until 2030. The resolution taken by the UN General Assembly 25 September 2015, A/RES/70/, 44 p.
- [3] Report of the interdepartmental group of experts for indicators of objective fulfillment in the sphere of sustainable development. UNO. Economic and Social Council. Forty-Seventh session. 8-11 March 2016. E/CN.3/2016/2, 46 p.
- [4] Khovanov N.V. Analysis and synthesis of indicators in the information deficit. SPb.: SPbSU Publishers, 1996. 196 p.
- [5] Noam Lior. Quantitative indicators of sustainable energy development. Energy Bulletin, 2015, No.19, p.8-29.
- [6] Dmitriev V.V., Osipov G.K. INTEGRAL ASSESSMENT OF STABILITY OF SOCIAL-ECOLOGICAL-ECONOMIC SYSTEM AGAINST CHANGES IN ITS FUNCTIONING CONDITIONS. 17th International multidisciplinary Scientific Geoconference SGEM 2017. Conference Proceedings. Volume 17. Ecology, Economics, Education and Legislation. ISSUE 52. Ecology and Environmental protection. 29 June – 5 July, 2017.Albena, Bulgaria, 565-572 pp. DOI: 10.5593/sgem2017/52.

[7] Dmitriev V.V., Lobacheva S.V., Chistilina V.S. Assessment of the state of socio-ecological-economic systems of the regions of Russia: AZR regions. Materials of the International scientific and practical conference "Modern Ecology: Education, Science, Practice", 2017, Voronezh, p.297-301.

[8] Dmitriev V.V. & Kaledin N.V. Integrated assessment of regional socio-ecological-economic systems and the quality of life (case study the constituent entities of the North-Western Federal District of Russia), Vol. 8, No.2, The Baltic region, Russia, 2016, pp.125-140.

[9] Vasiliy Dmitriev, Vladimir Kaledin, Nikolai Kaledin
THEORY AND PRACTICE OF INTEGRATED ASSESSMENT

OF THE STATE OF COMPLEX SYSTEMS IN NATURE AND SOCIETY 16th International multidisciplinary Scientific Geocoference SGEM 2016, Book 5, Ecology, Economics, Education and Legislation, www.sgem.org, SGEM2016 Conference Proceedings, ISBN 978-619-7105-66-7; ISSN 1314-2704, 30 June - 06 July, 2016, Albena, Bulgaria, Vol. II, 871-877 pp. DOI: 10.5593/SGEM2016/B52/S20.113.

[10] Dmitriev V.V. & Kaledin N.V. Integrated assessment of regional socio-ecological-economic systems and the quality of life (case study the constituent entities of the North-Western Federal District of Russia), Vol. 8, No.2, The Baltic region, Russia, 2016, pp.125-140

TRENDS IN DATA ANALYSIS: STATE, DEVELOPMENT PROSPECTS

Doctor in Economic sciences, Prof., I. A. Katsko; PhD in Economic sciences, P. Yu. Velichko;
Student, M. Nikogda
Kuban State Agrarian University – Russia, e-mail: stat@kubsau.ru

On the banks of the Rhine for many centuries was towering beautiful castle. The spiders, which dwelt in the cellars of the castle, tightened all its aisles with cobwebs. Once a strong gust of wind destroyed the thinnest threads of the web, and the spiders began to recover the gaps: they believed that the lock was kept on their web!

M. Klain [10]

Abstract. The article suggests the consideration of data analysis ideology in the context of knowledge creation process and the technological patterns of social development. The problems of singularity (human misunderstanding of data processing results) associated with increase in data variety, volume and further intellectualization of the corresponding technologies for their processing are proposed to be solved by creating new formalization techniques that allow retransmission.

KEY WORDS: DATA ANALYSIS, TECHNOLOGICAL PARADIGM, ANALYTICS, KNOWLEDGE, MACHINE LEARNING, BIG DATA AND THE INTERNET OF THINGS, MEGATRENDS, AMICABLE INTELLIGENCE OF THE HUMAN LEVEL, FORMALIZATION, CONSTRUCT, SCIP

1. Introduction. Data analysis (applied statistics) as development of statistics ideology, probability theory and mathematical statistics intensively developed over the last two centuries is naturally considered in the context of the socio-economic development of society with regard to solution of management and decision-making problems [16-18]. In this case, it seems interesting as a context to lead the ideology of technological paradigm¹, suggested by D.S. Lvov, S.Yu. Glazyev, G.G. Fetisov and essentially relied on larger cycles by N.D. Kondratiev (the phase of new ideas emergence - lasts about 10 years, the phase of paradigm growth - does about 40 years, the maturity phase lasts about 10 years more) [4, 5, 12]. Using the data analysis as an example, it is easy enough to trace the tendencies of the decision-making support ideology that are based on empirical observations characteristic for one or another way (Table 1)

2. Data analysis – contextual approach.

In postindustrial society, the cognitive revolution, which began in the 1950s and 1960s, manifested itself. They can talk about two periods of its development. I cognitive revolution, where a person is a carrier and a generator of knowledge, and a computer, the Internet and software are tools based on the machine learning ideology (the artificial intelligence implementation in a weak version - machine intelligence). II cognitive revolution - new knowledge is generated by the computer (the artificial intelligence formation).

The first scientific paradigms had a material basis, which is very important for the social adaptation of man in the real world. Cognitive paradigm, based on machine intelligence (and in the long term amicable *artificial intelligence of the human level* - AIHL (DIYCH-in Russian), called the fourth industrial revolution, is focused on accelerating all processes by integration at the expense of information technology and the Internet of all things, the basic megatrends of modern society (physical, digital and biological) [1, 19, 22-24].

The purpose of data analysis is to obtain new knowledge about the studied system using observations or differently to convolve (compress) existing information for solving applied problems of

analysis and explaining the features of the studied system functioning, management, forecasting (predication) and decision-making.

The main difference between applied statistics (data analysis) from mathematical statistics is the consideration of not only probabilistic but also geometric and logical nature of data, as well as the obtaining of convolutions by both formal algorithmic methods (classical methods of *multidimensional statistical analysis* - MSA) and not formal ones (machine learning and adapted methods of MSA).

According to E. Toffler's studies, nowadays, power in society is based on three basic elements: strength, money and knowledge. [21] Moreover, knowledge becomes a universal tool that can replace all others. This is why Russell Ackoff's definition, which characterizes the process of the knowledge formation, acquires a special meaning, which in our formulation is expressed in the following way [2]

Facts – Information – Data – Knowledge – Understanding – Wisdom.

One of the forms of knowledge representation contributing to thinking formation and worldview of human has always been mathematics, which allowed to form a chain of thinking levels (*recognition - reproduction of model situations - atypical situations analysis - creativity*). The most important stage in the process of knowledge formation is understanding - a person easily perceives and uses in practice what is understandable. From the point of view of data analysis, the stages of the knowledge formation can be disclosed in the following way [2, 6]:

- *facts – events, that have already happened;*
- *information – facts characteristic;*
- *data – facts, described quantitatively or qualitatively, presented in the form of tables «object – property (feature)» or «question - answer»;*
- *knowledge – rules «If ..., so ...», which can be used in decision-making;*
- *understanding – presentation about functional features of studied object, managing possibilities, foresight (predication) and decision-making;*
- *wisdom – ability to use the reached understanding in future.*

Table 1 – Data analysis in context of socio-economical development

Technological way	Main formalization form (approach)	Data processing way
I mechanization 1770-1830	Mathematical analysis (data – realization results of mathematical laws of natural sciences)	Descriptive statistics, differential and integral calculus
II steam engines, railways 1830-1880		
III electricity, metallurgy 1880-1930		
IV oil, mass production, nuclear power 1930-1980	Probability theory and mathematical statistics (data – random processes realization, which submit to certain distribution laws – parametrical statistics, or nonparametrical statistics)	Selective method, Convolution of information. Formalistic algorithmic approaches, solving problems: Data description, visualisation, classifications and dimension decrease, search of dependences (multidimensional statistical analysis)
V informatization, telecommunication 1980-2020	Data analysis (applied statistics) (any data nature: probabilistic, geometric, - data form in multidimensional attribute space «compact» (clots), logical – this not only quantitative, but also non-numerical (qualitative) form patterns – interrelations not always explainable at the quantitative level)	Analytics 1.0 (descriptive analytics) OLAP cubes, convolution procedures that do not allow an algorithmic approach - Exploratory data analysis (EDA), also based on the computer training ideology (Data Mining) as an option for implementing EDA based on information technology, web-sites scraping
VI nano-, bio-, info-, cognitive-, socio-technology – NBICS 2020-2060	<i>Big Data, Internet of Things (IoT)</i> (data nature is any of the above, including visual, textual, sound, video-audiofiles and others)	Analytics 2.0 (predictive analytics) Analytics 3.0 (prescriptive analytics - the basis of CRM) Analytics N.0 (analytics, supporting typical solutions, having opportunities to search and process information on-line/interface (analogue of modern app Siri from Apple))
VII human – main technology subject 2060-2100	... Data nature is any of the above, including not representable in semiotics systems projection foresight	Analytics NBICS, based on technologies of amicable intelligence of human level

In order to use methods of applied statistics, the data must be measured with qualitative or quantitative scales. The measurement process is accompanied by problems: heterogeneity, quality,

limitations, subjectivity of perception and thinking. Moreover, person is limited in perception of the surrounding world. As it is known, it is characterized by the number of J. Miller (1956) (7 ± 2) - according to it, there is a need to compress large volumes of information and representation in the form of (preferably understandable) models. This goal is devoted to the work of decision support systems that allow solving tasks: descriptive statistics (OLAP cubes), classification and diminution of dimensions, search for dependencies, prediction, etc. implemented in KDD class systems and Data Mining.

The implementation of machine learning methods does not allow us to realize the "understanding" stage in R.Akoff's knowledge formation process and shows the practical application of Braimean's uncertainty principle, which is an analogue of Heisenberg's uncertainty principle in data analysis context:

$$\begin{aligned} &\text{«accuracy} \times \text{interpretability} = \\ &= \text{Braimean's constant} \end{aligned}$$

The famous futurist E. Toffler talks about three waves in the development of society: agrarian, industrial, informational, which their traditional education systems conformed to [19]. For several decades, we have witnessed the transformation of education traditional system for an industrial society. Data analysis has always evolved in the direction of meeting the needs of society. If there was enough descriptive statistics in the agrarian society, in industrial - analytical statistics, the postindustrial society and the expected information extended the applied statistics with machine learning methods using both structured and unstructured data (Data Mining, Text Mining, Web Mining, Social Mining, Big Data and the Internet of things), thus, the demand for data analysis methods is determined by the level of development of the society, for example, business intelligence technologies that have long been used in Moscow and St. Petersburg, are developing in the regions only in recent years, becoming one of the costly business articles (Table 1). The development of the digital sector of the economy is being discussed in the world, therefore, it becomes necessary to systematically comprehend the possibilities, limitations and prospects for data analysis at the present stage of society development.

3. Megatrends. Today's data processing methods are based on the ideology of probability, statistics, data analysis and, among other tasks, allow solving the problem of finding "megatrends" (conditional coordinate system) at different levels of society at a new level that allows explaining many phenomena in the socio-economic space. The number of "megatrends" for a society is the same as for a person (7 ± 2) and corresponds to one of the theories claiming the theory of "great unification" in physics - superstring theory, which requires about 10 measurements. At present, one of the obvious "megatrends" explaining the transformation of the education system is that "the transition to an information society requires getting rid of the outdated educational system that trained cadres for the industrial society." Analysis of printed sources allows us today to talk about the following "megatrends" of our society, conditioned by the digital revolution and carrying a disruptive influence (a form of natural selection in biology), "tearing" the homogeneous aggregate (usually) into two extreme variants and not contributing to the average state, which explains the appearance in the socio-economic systems of power law distribution laws, such as the Pareto or Zipf law (for example, the result of such an impact can be considered the stratification of society: rich and poor, "golden" Billion and others, etc.) [1, 2, 8, 9, 19-24] :

- 1) Mankind today lives in a consumer society and gradually loses the distinction between the real and the virtual.
- 2) Priority in life is obtained by "physical" people, without moral and moral obligations.
- 3) The society increasingly depends on information technology (3D-printing, development of 4D-printing

technology, the production of any services and goods in online mode).

- 4) Information technology is becoming one of the subjects of everyday life of modern man (unmanned vehicles, robotics, new materials).
- 5) People pay and receive money for virtual actions that do not give a sense of physical incarnation, which negatively affects the human psyche.
- 6) New approaches to interaction and cooperation at all levels of the society are being developed. For example, "distributed databases" are *block chains* that represent a data store that is available for verification to anyone (for example, *Bitcoin*).
- 7) Particular attention is paid to *Big Data* concepts and the "*Internet of all things*" in the study of social networks, industry, business.
- 8) The growing opportunities for biological engineering require the development of a normative, ethical and legal framework.
- 9) The transition to an information society requires transforming an outdated education system that trained personnel for an industrial society by selecting a goal at the state level (for example, harmonious development of a person) instead of replenishing a certain labor market (or "human cloud").
- 10) The actual sector of the economy, based on the "human cloud", is becoming relevant.

Modern information technologies for data analysis (*web mining* and *text mining*, etc.) make it possible to find an alternative to classical content analysis when searching for "megatrends" including without human participation (scraping web-sites).

"Megatrends" change over time and the past ("megatrends") "twists" along the new ones, relying on the general direction (mainstream) of the 21st century - a digital revolution that, like the communist movement 100 years ago, will change the landscape of the planet. Perhaps right now there is a tectonic gap between the "golden billion" and the rest of the world's population, although the maturing changes are not universal even for developed countries. For example, the politics of the consumer society is unacceptable in the Arab world and can have a different form (as in India and China). In Russia there is a centro-peripheral model of socio-economic space (N. Zubarevich) [8]:

- post industrial Russia (federal cities with a million population with postindustrial economy),
- industrial Russia (industrial cities with a population of up to 250 thousand people),
- rural periphery
- agrarian Russia (the main part of the country and residents of settlements with a population of less than 20 thousand people) ,
- patriarchal republics, based on their own values (the North Caucasus, Southern Siberia).

And all four of Russia perceive different future changes and react differently.

At different levels of the hierarchy of society, there are their own tendencies to change the world, so at the state level (in most countries) first of all one can distinguish: bitcoin, crypto-currencies, cyberwar, fakes.

4. The problem of translation of knowledge. Current trends in the development of analytics are directly related to the achievements of human intellect, which cause a futurist (fear of the future). Mankind is preoccupied with its own ideas about the power of computers and the alleged consequences of the emergence of artificial intelligence (AI). Many scientists expect the emergence of the point of "singularity" - the moment when the possibilities and results of the activity of artificial intelligence systems in the narrow sense (understood as the realization of the ideology of machine learning) will surpass the possibilities of human understanding [1,

9]. In addition, artificial intelligence is expected to reach the human level. The only thing people hope for is that they will be a *friendly human intellect (AIHL)*.

Intelligence of information systems is achieved today through the use of the Internet (the Internet of things), the ideology of machine learning, and computational capabilities. Thus, we are not talking about the presence of consciousness, but it is possible that the opportunity will come of "creating the effect of consciousness" through the effect of replacing the computational abilities. Then only the initial education of unconditional friendliness to man will pass the point of singularity.

In fact, we are talking about machine intelligence, which, due to the processing power, can generate new knowledge in the form of patterns and (or) constructs that can not be explained by humans on the basis of available information (including using *Big Data* and the Internet of all things). The traditional sequence of the process of forming knowledge according to R. Akoff (Facts - Information - Data - Knowledge - Understanding - Wisdom) will be broken. For if traditionally for a person of "knowledge" are products (rules) "if ..., then ..." that allow to realize the stage of "understanding", then the question arises about the need for a new round of development of mathematics and information technologies oriented to "not ... exclusion "of man from the processes of management and decision-making in the socio-economic space, since the knowledge obtained by the *AIHL* can go beyond the boundaries of human understanding (the rules" if ..., then ... "). Data analysis today is a way of compressing large volumes of information to support decision-making processes, which allows to identify patterns in data and to present them in the form of: graphs, tables, formulas, various dependencies obtained using machine learning methods. It is assumed that at the point of singularity, knowledge will go beyond these limits.

The problem of understanding the intellectual systems of the future is similar to the problems of the middle of the last century, when the possibilities of interaction with computers were formed through programming languages, which today number more than 8000.

We believe that in order to solve the problem of "retransmission" of knowledge to a person, the potential capabilities of the *AIHL* must presuppose the possibility of synthesizing a number of subject areas (SA) into which new knowledge can be projected. The lexicographic ordering of the SA will allow to identify and rank the consequences of applying new knowledge in different areas.

Thus, it becomes urgent to develop an understandable ideology of the description of the SA. As a possible example, consider, following the work of L.S. Bolotova [3], the method of situational analysis and design of the design of the SA model, on the basis of the set-theoretical (relational) approach, in the form of a complex of invariant constructs as applied to the description of the SA for the new knowledge.

4.1. The domain model. The synthesized object (system) should be created under the condition of the existence of an external environment that is characterized by a certain subject area (SA) - a part of the real world within the given context (industrial, agricultural, financial, computer, etc. corresponding to the direction of knowledge). Each subject area has its own language, which can be formalized using binary relations [3]. Usually, a system is understood as the set of a related set of objects. Most often, two sides of connectivity are considered: as a fact of the existence of a relationship between individual elements of the system - realizing the cognitive conceptual aspect (cognitive maps); as a description of the process of the corresponding connectivity of elements - a functional, information or behavioural aspect (semantic networks, frames, products, methods of situational modelling). Both approaches are considered rather rarely.

Modelling SA of an arbitrary nature is connected, first of all, with the analysis of the categories describing it. A category is understood as a construct or otherwise some abstract container, with some objects entering into it, and others not (this is the postulate of

our thinking for more than two millennia). It is assumed that the categories that a person operates on can be arranged in the following hierarchy: the higher level - the base level - the lower level. The base level is the level at which most of our knowledge is structured. It gives an opportunity to perceive geometric visualization of the conceptual structure of an object.

4.2. Algebraic models of SA. At present, algebraic language and style of thinking are the standard approach to the representation of data and knowledge in information systems. The question of the possibility of a correct description of the model of the subject domain associated with the person (observer) in the form of a system of objects with certain relations can be investigated only by means external to this system (that is, in some other theory), as follows from K. Godel's theorem on incompleteness of formal arithmetic (to which almost all mathematical theories can be reduced).

D. Hilbert describes the interpretation of the formalization of mathematical theory, as well as the method that makes the formal system the subject of the study of mathematical discipline - metamathematics or the theory of evidence [11].

The introduction of a system of S objects according to the ideology of metamathematics, assuming the existence of a non-empty set of objects between which certain relations are established, can proceed from two methods characterizing two main trends in modern mathematics - constructivism (modern predecessor of which is A. Poincaré and which was basic in ancient science, for example, in Euclid) and formalism, which implies a complete abstraction from the meaning.

In an axiomatic method, the axioms underlying the formal approach are used as assumptions about the system of S objects. Then we examine the consequences of the axioms, which form a theory with respect to the system of S objects under consideration.

A constructive (genetic) method involves constructing objects in a certain order. S. Klini characterized this approach as a method of substantive or material axiomatics. To describe the system from the point of view of the observer observing the system from the outside, a "formal system object" is introduced-the meta-set of the research system S , which allows describing systems based on logical and mathematical methods [13].

The presence of experts with exact, technological, effective thinking, free "from traditions and cognitive prejudices", which will interact with AIHL, is postulated. To do this, special (psychological) work with experts and subjects of the problem of creating a new object is supposed, accustoming them to operate with their "constructs".

The basic representation of the construct is the meta-set

$$S < X_{as}, X_a, X_{ao}, \cup X_{ac_i} >,$$

where X_{as} – (action subject), X_a – (action), X_{ao} – (action object), $\cup X_{ac_i}$ – (action components). All elements of meta-set have (property):

$X_{as} = X_{as}(p_{s1}, \dots, p_{sl})$, $X_a = X_a(p_{a1}, \dots, p_{aq})$, $X_{ao} = X_{ao}(p_{o1}, \dots, p_{ot})$, $X_{ac_i} = X_{ac_i}(p_{ac_i1}, \dots, p_{ac_ih})$, the results of relations of which among themselves within the framework of our task realize TER (technological, technical, operational, economic, environmental requirements for the subject area, which are based on the normative provisions defined by the person).

A logical representation of a construct implies two components:

1) functional – Φ , regarded to the purposes of building a new object and described as a union of binary relations (R)

$$\Phi = R_{as}(X_{as}, X_a) \cup R_{ao}(X_a, X_{ao});$$

2) providing, (achievement of the goal) – Q ,

$$Q = R_{as}(X_{as}, X_a) \cup R_{ac_i}(X_a, X_{ac_i}).$$

$$K = \Phi \cup Q.$$

Each construct K is a kind of domain concept that is open to expansion and modification, which is intended for reusable use in designing, obtaining production rules, and so on.

Combining all constructs:

$$U = \cup K_j$$

gives us the universum U , a structure called a polyhedron in topology, describing an SA of arbitrary nature.

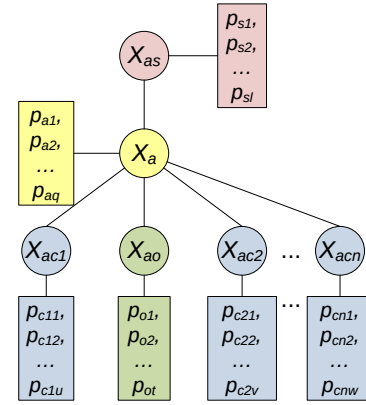


Figure 1 – Construct model – Subject area extract [3]

The universum is a generalized model of a specific subject area, which can be represented as a basis for concepts (ontologies) designed for reusable, multipurpose use in various applications and relationships between them that implement production rules [14]. Consideration of ontologies with selection functions and mechanisms for their implementation allows us to talk about a knowledge base that potentially allows the formation of products (rules) that are understandable to man [15].

5. Conclusions. The problems of human interaction and the human-level friendly intellect require the creation of means for "retransmitting" to a person new knowledge obtained by advanced methods of data analysis (structured, weakly structured, unstructured). At the author's level of vision, developments are required that allow translating, obtained knowledge (patterns, constructs, etc.) into an "understandable" kind of person, for example, projecting into several mutually complementary subject areas, which can be realized (AIHL) described by the method of situational analysis and design of the design of the SA model. Within the framework of the domain model - the creation of a new object (or description of the situation) is reduced to operations over the relations between constructs and their elements. Thus, one of the variants of "passing the point of singularity" is the formation of knowledge bases related to a certain subject area and an assessment of the consequences of the implementation of new knowledge through the use of scenarios based on the SA model.

The emergence of a new scientific paradigm in science and society forms a new space in which the previous includes (due to) a new generalization, the postulates (principles) change and shrink, with an increase in the coverage of the phenomena described (the scaling effect of socio-economic space) [7], which is realized in an explicit form using the example of data analysis. The further development of intelligent information systems leads to the automation of the work of analysts and other professionals, the emergence of new knowledge and the ability to present them for adequate human perception, and the need to create ethics councils, to address employment and social issues is already generally recognized. To preserve one's identity in the future information society, a person needs to solve many problems: preserve human culture, universal values, find for himself and implement new means of formalization (adapt or develop new mathematics for translation of knowledge obtained by the AIHL), etc. Everyone should understand that it is he who builds the future world and how he will solve it.

6. Literature

1. J. Barrat, The last invention of mankind: Artificial intellect and the end of the era of *Homo sapiens*. /from English. Natalia Lisova. - M.: Alpina non-fiction, 2015. - 304 p.
2. P.S. Bondarenko Theory of Probability and Mathematical Statistics: Textbook / P.S. Bondarenko, G.V. Gorelova, I.A. Katsko; Ed. by I.A. Katsko, A.I. Trubilina. Moscow: KNORUS, 2017. - 390 p.

3. L.S. Bolotova. Conceptual design of the domain model with the help of software systems for the development of knowledge bases for intelligent decision support systems / L.S. Bolotova, V.A. Smolyaninova, S.S. Smirnov // High technology: scientific. -technical. - 2009. - T. 10 No. 8. - P. 28-36.
4. S. Yu. Glazyev, D.S. Lvov, G.G. Fetisov. Evolution of technical and economic systems: the possibilities and boundaries of centralized regulation. - M.: Science. - 1992. - 208 p.
5. S.Yu. Glazyev. Strategy of advanced development of Russia in the conditions of global crisis. - M.: Economics, 2010. - 255 p.
6. N.G. Zagoruyko. Applied methods of data analysis and knowledge. - Novosibirsk, Uni of maths, 1999. - 270 p.
7. V. A. Zorich Mathematical analysis of natural science tasks - M.: MCNMO, 2017. - 160 p.
8. N. Zubarevich. «Four Russia-s» and new political reality. Web: http://polit.ru/article/2016/01/17/four_russians/
9. M. Kaku, Mind future / from eng. by N. Lisova, - M.: Alpina non-fiction, 2015. - 502 p.
10. M. Kline. Mathematics. Loss of certainty. - The World, 1984. - 134 p.
11. S. Klini. Introduction to metamathematics: - M.: Librocom, 2009. - 528 p.
12. N.D. Kondratiev, Large cycles of conjuncture and theory of foresight / edited by AL Albakina. - M.: "Publishing house" Economics ", 2002. - 767 p.
13. V.V. Kulba etc. Information processes and information management / Human factor in management c6. Articles of the ISP RAS ed. O.N. Abramova, K.S. Ginsberg, D.A. Novikov and others - M.: KomKniga, 2006. - 496 p.
14. Y. Lyapar. Theory of System-Structural Design - the Basis of Intellectualization of Modelling and Decision Support Systems / Yu.I. Lyapar, I.A. Katsko, G.F. Bershitskaya - Krasnodar.: KSAU, 2010. - 49 p.
15. Yu. I. Lyapar. Synthesis of knowledge bases of analog electronic devices. // Ulyanovsk: Works of the Intern. Conf. "Continual and Algebraic Logic, Calculus and Neuroinformatics in Science and Technology", vol. 3, 2006, p.120-128.
16. A.V.Maltseva, N.E. Shilkina, O.V. Mahnitkina Data mining sociology: Experience and outlook for research // Social surveys 2016-January(3), p. 35-44.
17. D.A. Novikov. Cybernetics: Navigator. History of cybernetics, current state, development prospects. - M.: LENAND, 2016. - 160 p. (Smart Management Series)
18. A.I. Orlov. The main features of the new paradigm of mathematical statistics // Polytematic network electronic scientific journal of the Kuban State Agrarian University (Kuban State University, Krasnodar, KubSAU, 2013. № 06 (090), P. 188-214 <http://ej.kubagro.ru/>
19. D. Rose, The Future of Things: How a Fairy Tale and Fantasy Become a Reality / 3rd edition / translation from the English Seeds of Sheshenin. - M.: Alpina non-fiction, 2017. - 344 pp.
20. E. Toffler. The Third Wave - M.: AST, 2010. - 784 pp.
21. E. Toffler. Metamorphoses of Power - M.: AST, 2004. - 672 pp.
22. M. Ford. Technologies, who will change the world. Alexandra Kardash, M.: Mann, Ivanov and Ferber, 2014. - 268 p.
23. M. Ford. Robots come: Technology development and the future without work / Translated from English by Sergey Chernin. - M.: Alpina non-fiction, 2016. - 430 p.
24. K. Schwab. The Fourth Industrial Revolution / K. Schwab - M.: "E", 2017. - 208 p.

EXACT RECONSTRUCTION VERSION OF RADON TRANSFORMATION IN TOMOSYNTHESIS

Morgun O., Nemchenko K., Vaisburd A.*, Viktynska T*.
 "Radioprom" LTD, Dostoevskogo Str. 1, Kharkov, Ukraine 61011
 *Karazin Kharkov National University, Svobody Sq.4, Ukraine 61022

Abstract: The purpose of this study is a method of exact reconstruction in the Radon problem, which consists in refusal of the approximate transformation kernel usage. A comparison of the methods that are currently used in tomosynthesis was conducted. Model experiments were performed; the results of the application of the proposed method in real tomography studies in tomography are given.

The traditional methods of Radon transformation [1] do not make it possible to accurately reconstruct the initial distribution function due to the divergence of the integral expression. In this

paper, we suggest alternative method aiming to get rid of this defect by changing the order of integration using re-grouping the integrand functions in the result of transformation:

$$f(x, y) = \int \frac{|Q|dQd\alpha dX}{(2\pi)} F(X, \alpha) \exp(-iQ(X - x \sin \alpha + y \cos \alpha)). \quad (1)$$

Here $f(x, y)$ is the desired density distribution in the layer, $F(X, \alpha)$ is the projections made under angle α , and X is the coordinate in the detector. Usual way to use (1) is introduce function (filter)

$$G(z) = \int \frac{|Q|dQ}{(2\pi)} \exp(-iQz) = \frac{1}{\pi} \int_0^\infty |Q| \cos(Qz) dz, \quad (2)$$

which describes the connection between $f(x, y)$ and $F(X, \alpha)$:

$$f(x, y) = \int d\alpha dX F(X, \alpha) G(X - x \sin \alpha + y \cos \alpha). \quad (3)$$

The main problem here is that the integral (2) contains a singularity and, therefore, approximate functions [2] are used instead of it, which reflect the main features of the function (2), but lead to an approximate recovery.

In this paper, we propose to change the order of integration to avoid this problem. In this case none of the integrals into this

expression (1) diverges within the limits. As a result of taking the integral in this way, we eventually have to obtain the distribution function equivalent to the original one.

Let's consider the case where the initial distribution function is the delta function, which is determined by the following expression

$$f(x, y) = \rho_0 a^2 \delta(x - x_0)(y - y_0) \quad (4)$$

which corresponds to the projection:

$$F(X, \alpha) = \rho_0 a^2 \delta(x_0 \sin(\alpha) - y_0 \cos(\alpha) - X) \quad (5)$$

Whatever divergences arise in the integral expression (1), we shall not get it to the form containing the kernel of the Radon transformations (2) in explicit form, but immediately substitute the expression (5) into it. Thus, the relation (1), taking into account the integration limits, can be written in the form:

$$f(x, y) = \frac{\rho_0 a^2}{(2\pi)^2} \int_{-\infty}^{+\infty} Q dQ \int_0^\pi d\alpha \int_{-\infty}^{+\infty} dX \delta(x_0 \sin(\alpha) - y_0 \cos(\alpha) - X) * e^{-iQX} e^{iQ(x \sin(\alpha) - y \cos(\alpha))} \quad (6)$$

Now integrating

this expression first with dX , getting rid of the delta functions. As a result, we get:

$$f(x, y) = \frac{\rho_0 a^2}{(2\pi)^2} \int_{-\infty}^{+\infty} Q dQ \int_0^\pi d\alpha e^{iQ((x-x_0) \sin(\alpha) - (y-y_0) \cos(\alpha))} \quad (7)$$

Now we return to other integration variables to avoid divergence and make integration possible in a different order. It is logical to go over to the variables $(\vec{k}_x, \vec{k}_y)$, for this we write the Jacobian of the transition:

$$dQd\alpha = \left| \begin{matrix} \widetilde{k}_x = Q \sin(\alpha) \\ \widetilde{k}_y = Q \cos(\alpha) \end{matrix} \right| = d\tilde{k}_x d\tilde{k}_y \frac{\partial(Q, \alpha)}{\partial(k_x, k_y)} = \frac{d\tilde{k}_x d\tilde{k}_y}{Q}, \quad (8)$$

taking this into account, expression (7) can be written as follows:

$$f(x, y) = \frac{\rho_0 a^2}{(2\pi)^2} \int_{-\infty}^{+\infty} e^{i\tilde{k}_x(x-x_0)} d\tilde{k}_x \int_0^\infty e^{-i\tilde{k}_y(y-y_0)} d\tilde{k}_y. \quad (9)$$

It is easy to see that the two integrals in the expression are just the Fourier representation of the delta function, similar to the representation:

$$\delta(x - x_1) = \int \frac{dQ}{2\pi} e^{-iQ(x-x_1)} \quad (10)$$

Thus, we get:

$$f(x, y) = \frac{\rho_0 a^2}{(2\pi)^2} 2\pi \delta(x - x_0) 2\pi \delta(y - y_0) \quad (11)$$

Hence, we get the answer:

$$(x, y) = \rho_0 a^2 \delta(x - x_0) \delta(y - y_0) \quad (12)$$

We can see that this function completely identical to the function (4) introduced by us, moreover, we had no need to enter redefinitions anywhere, since all the transitions were initially equal. As a result, we showed that the delta distribution function can be accurately reconstructed by the proposed method.

After the prove above that the delta function can be accurately reconstructed, natural to assume that any other function can also be reconstructed without loss of precision. We can prove it

by taking into account that any distribution function can be represented as a continuous set of delta functions, and all the resulting interim expressions are additive quantities, which gives the right to put the sum or integral sign before the reconstructed distribution function.

This explanation can be considered as self-explanatory. Nevertheless, we strictly prove this statement without referring to the result already obtained. To do this, we return to the previously obtained formula (1) and change the order of integration:

$$f(x, y) = \frac{1}{(2\pi)^2} \int_0^\infty |Q| dQ \int_0^\pi d\alpha \int_{-\infty}^{+\infty} e^{-iQX} e^{iQx \sin(\alpha) - iQy \cos(\alpha)} F(X, \alpha) dX, \quad (13)$$

now the integration over dQ will be performed last.

To converge this integral, we recall the linear distribution

$$F(X, \alpha) = \int f(x, y) dl, \quad (14)$$

and represent

$f(x, y)$ through the delta function as follows:

$$f(x, y) = \iint d\tilde{x} d\tilde{y} \delta(x - \tilde{x}) \delta(y - \tilde{y}) f(\tilde{x}, \tilde{y}), \quad (15)$$

Then we use the following presentations

$$F(X, \alpha) = \int dx dy \delta(x \sin(\alpha) - y \cos(\alpha) - X) f(x, y) \quad (16)$$

$$dl = dx dy \delta(x \sin(\alpha) - y \cos(\alpha) - X) \quad (17)$$

Putting this into (13) we get:

$$(x, y) = \frac{1}{(2\pi)^2} \int_0^\infty |Q| dQ \int_0^\pi d\alpha \int_{-\infty}^{+\infty} dX e^{-iQX} e^{iQx \sin(\alpha) - iQy \cos(\alpha)} \iiint d\tilde{x} d\tilde{y} \delta(x - \tilde{x}) \delta(y - \tilde{y}) f(\tilde{x}, \tilde{y}) dx dy \delta(x \sin(\alpha) - y \cos(\alpha) - X) \quad (18)$$

Integrating by $d\alpha$, we get:

$$f(x, y) = \frac{1}{(2\pi)^2} \int_0^\infty |Q| dQ \int_0^\pi d\alpha \int_{-\infty}^{+\infty} dX e^{-iQX} e^{iQx \sin(\alpha) - iQy \cos(\alpha)} \iint d\tilde{x} d\tilde{y} f(\tilde{x}, \tilde{y}) \delta(\tilde{x} \sin(\alpha) - \tilde{y} \cos(\alpha) - X), \quad (19)$$

then integrate over dX and get:

$$f(x, y) = \frac{1}{(2\pi)^2} \int_0^\infty |Q| dQ \int_0^\pi d\alpha e^{iQx \sin(\alpha) - iQy \cos(\alpha)} \iint e^{-iQ\tilde{x} \sin(\alpha) + iQ\tilde{y} \cos(\alpha)} f(\tilde{x}, \tilde{y}) d\tilde{x} d\tilde{y} \quad (20)$$

Now we make the transition to new variables completely analogous to the transition (8), so we easily receive from (7) the expression:

$$f(x, y) = \frac{1}{(2\pi)^2} \int dk_x dk_y e^{ik_x(x-\tilde{x})} e^{ik_y(y-\tilde{y})} \iint f(\tilde{x}, \tilde{y}) d\tilde{x} d\tilde{y} \quad (30)$$

Let's consider the representation of the delta function again:

$$\delta(x - x_1) = \int \frac{dQ}{2\pi} e^{-iQ(x-x_1)} \quad (31)$$

Applying it we get:

$$f(x, y) = \iint d\tilde{x} d\tilde{y} \delta(x - \tilde{x}) \delta(y - \tilde{y}) f(\tilde{x}, \tilde{y}) \quad (32)$$

Comparing this result, we can see that this expression is fully equivalent to the expression (15).

Finally, we conclude that the proposed method absolutely allows reconstructing the initial distribution, regardless of what function it describes. Let us highlight, that this proof doesn't contain any reference to the result of the reconstruction of the delta-like distribution function, so the proof done in the previous section can be considered as a particular case.

REFERENCES

1. Radon, Johann (1917), "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser

Mannigfaltigkeiten", *Berichte über die Verhandlungen der Königlich-Sächsischen Akademie der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse [Reports on the proceedings of the Royal Saxonian Academy of Sciences at Leipzig, mathematical and physical section]*, Leipzig: Teubner (69): 262–277; Translation: Radon, J.; Parks, P.C. (translator) (1986), "On the determination of functions from their integral values along certain manifolds", *IEEE Transactions on Medical Imaging*, 5 (4): 170–176, PMID 18244009, doi:10.1109/TMI.1986.4307775.

2. Herman, Gabor T. (2009), *Fundamentals of Computerized Tomography: Image Reconstruction from Projections* (2nd ed.), Springer, ISBN 978-1-85233-617-2

APPLICATION OF FUZZY MODELING TO PREDICT THE DISEASE OF STAFF FROM EXPOSURE TO WORKING CONDITIONS

Klimova I. V., Smirnov Yu. G.

Department of Informatics, Computer Technologies and Engineering Graphics - Ukhta State Technical University, the Russian Federation

bgd4@mail.ru, ysmirnov@ugtu.net

Abstract: A fuzzy model for determining the morbidity rate of employees of a refinery with diseases of the respiratory organs is analyzed on the basis of an analysis of the concentrations of pollutants in all occupational environments using a mathematical apparatus of fuzzy sets. The results of visualization of the developed fuzzy model in the MATLAB Fuzzy Logic Toolbox medium are presented.

Keywords: FUZZY MODELING, MODEL, WORKING CONDITIONS, HARMFUL PRODUCTION FACTORS, MORBIDITY, STAFF

1. Introduction

Fuzzy modeling is the most promising direction for scientific research in the field of analysis, forecasting and modeling of various processes. This is especially important for assessing occupational risks for the health of personnel, where there is insufficient data on the connection of certain diseases with working conditions. The existing assessment of working conditions makes it possible to determine the "verbal" level of risk on the basis of the established class of working conditions. At the same time, the range of production factors under consideration is constantly narrowing, excluding even the receipt of additional payments for harmful working conditions. Thus, with the invariance of working conditions, the class of working conditions is reduced only by changing the methodology by which labor conditions are evaluated.

The purpose of this work is to assess the applicability of existing models and methods of fuzzy logic to modeling the impact of harmful substances on the health of personnel of "RN - Komsomolsk Oil Refinery".

The object of the study is the personnel of the plant exposed to harmful substances in three environments for the period from 2004 to 2010. The subject of the study is inhalation non-carcinogenic risks for the health of the personnel of the oil refinery.

2. Problem discussion

Consider the problem of constructing the dependence of the morbidity of diseases of the respiratory organs of the personnel of the oil refinery on the indices of the non-carcinogenic hazard of chemicals and suspended substances. The nosological group of "respiratory diseases" was chosen on the basis of an earlier analysis of the impact substances and the peculiarities of their effect in three environments: industrial-technological, industrial and urban [1].

The production technological environment is the most polluted - it is the territory of technological workshops and installations. The production environment is a less polluted environment within administrative buildings, as well as the rest of the plant's territory. The urban environment in turn is a combination of household and environmental environments in the city of Komsomolsk-on-Amur. Such a division into environments is made with the aim of clarifying the concentrations of harmful substances, and, accordingly, a more detailed assessment of the health risks of workers [1].

A total of 66 substances participated in the analysis. Figure 1 shows the quantitative ratio of substances affecting personnel. Of them, substances affecting the respiratory system - 30 items, 7 has a primary effect on the respiratory system (Figure 2) [2].

All harmful substances that affect the respiratory system can be divided into two groups: due to their specific effects: chemical substances and suspended substances. Therefore, the model of the dependence of the morbidity of personnel with respiratory diseases on two parameters will be constructed: the index of non-

carcinogenic hazard for chemical substances and the index of non-carcinogenic hazard for suspended substances.

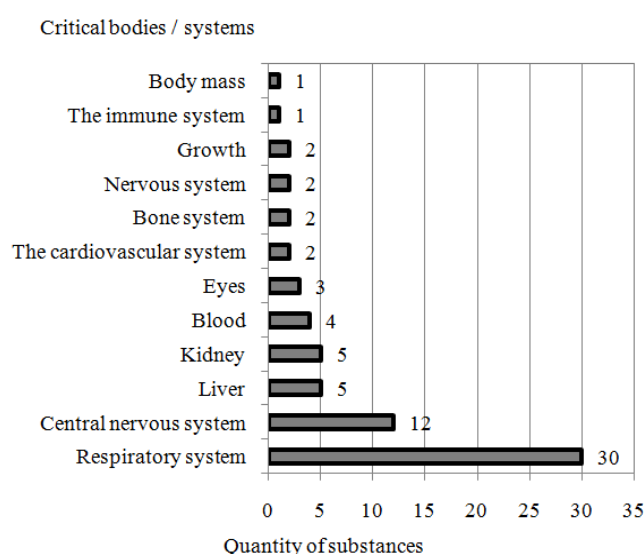


Fig. 1. Ratio of substances affecting critical bodies / systems

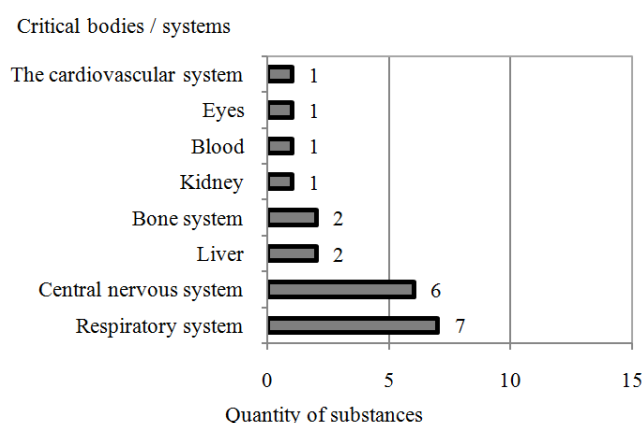


Fig. 2. Ratio of substances according to their primary effect on critical organs / systems

Indices of non-carcinogenic hazards are quantitative estimates of the amount of harmful substances affecting the worker's body throughout the day in three environments and are calculated first for each substance (1) and then for the group of effects on the particular organ or system as a whole (2):

$$(1) HQ_i = AC / RfC,$$

$$(2) HI = \sum HQ_i,$$

where HQ_i are the hazard ratios for the individual components of the mixture of agents; HI - index of non-carcinogenic hazard for a critical organ or system; AC - average concentration of substance, mg/m^3 ; RfC - reference (safe) concentration, mg/m^3 [1].

It is worth noting that the average concentration of each substance involved in the calculation was averaged over 24 hours, depending on each type of medium, the concentration of the substance in it, and the time of exposure to the worker.

After the HI indices for respiratory organs were calculated at each workplace, the data were averaged over 29 occupational groups. The main criteria for the formation of groups of personnel: belonging to the facility or shop and the value of the received non-carcinogenic risk HI .

The minimum values of the indices were $HI_{min\ chem} = 23.87$ and $HI_{min\ dust} = 2.56$, which exceeds the permissible value $HI_{norm} = 1$. The maximum values of the indices were $HI_{max\ chem} = 1621.58$ and $HI_{max\ dust} = 5.01$.

Next, consider such an output parameter as the state of health of personnel, which is a reflection of a complex set of phenomena in the environment. The process of its formation is influenced by a number of industrial, socio-economic, as well as biological, anthropogenic, natural climatic and other factors that together determine the ecological environment in which the person is during the day, food and water. Most of the harmful substances a person receives with inhaled air, this is about 80% of all intake doses.

Anticipate the corresponding incidence rate of personnel with respiratory diseases is quite difficult, especially with the seasonality of this phenomenon. In this connection, it is expedient to use linguistic variables, that is, variables whose values are not numbers, but words in natural or formal language.

The data on the morbidity of the personnel of the plant on the nosological form of "respiratory disease" from 2004 to 2010 were broken down into the allocated 29 occupational groups, and then averaged over 7 years and included the number of cases per 1000 workers and the number of days of incapacity for work 1000 working. Further, data are used only for the number of cases of diseases: the minimum value is 125 diseases, the maximum is 495.

Comparison of calculated HI indices and morbidity leads to the receipt of 29 points, which are difficult to describe with the help of the equations of dependence, but nevertheless, a directly proportional relationship is clearly visible: the greater the hazard index, the greater the response.

For the transition to fuzzy logic, an additional conversion of the input parameters (HI indices) was performed. Since all indices exceeded the allowable value, an attempt was made to select a so-called "acceptable" value that would guarantee minimal deviations in the state of health.

For this purpose, an additional criterion "severity of disease progression" was introduced and a corresponding group was chosen whose indices are accepted for an "acceptable" new standard. This made it possible to obtain new ranges of values of $HI_{chem} \in [0.12, 8.05]$ and $HI_{dust} \in [1.0, 1.96]$.

3. Methodology

The modeling process can be presented in large form in the following sequence of actions:

- 1) awareness of the problem;
- 2) highlight the main factors that determine the problem, which should serve as output parameters of the model;
- 3) highlighting the defining input variables of the model;
- 4) the actual development of a mathematical model;
- 5) identification of the model (parametric or structural-parametric);
- 6) conducting numerical experiments with the model and, if necessary, statistical processing of the obtained data;

7) determining the composition of the quality parameters that characterize the problem being solved (using simulation data and a priori information);

8) formalization of particular quality criteria based on quality parameters;

9) determination of parameters characterizing the relative importance of particular criteria for solving a common problem;

10) formalization of the generalized quality criterion for solving a problem on the basis of aggregation of particular criteria, taking into account their relative importance;

11) solving the problem of choosing the best alternative or multi-criteria optimization, depending on the type of problem being solved.

In essence, it is some detail of the generally accepted scheme: the formulation of the problem $\rightarrow$ model building and identification $\rightarrow$ optimization.

The development of fuzzy models of sanitary and toxicological safety of the personnel of industrial enterprises makes it possible to obtain a numerical estimate of occupational risk. The mathematical apparatus of fuzzy logic is usually used in those cases when the available quantitative information is insufficient, or it is not complete enough to obtain reliable statistically significant conclusions [3-7].

Along with classical analytical methods, it is advisable to use the fuzzy set device implemented in particular in the MATLAB computer simulation system [8], which allows developing a fuzzy-multiple model for the estimation, analysis and visualization of professional risk indicators.

Since there is a need to take into account the multitude of indicators that are dissimilar in physical nature and dimensions, it is advisable to bring them to a dimensionless form by rationing, for example, as follows:

$$(3) s = \frac{s_i - s_{min}}{s_{max} - s_{min}},$$

where s_i is a normed index; s_{max} , s_{min} - the maximum and minimum value of the criterion in the sample according to the normed indicator.

The system of fuzzy inference in the general case includes the following stages:

1. Phasing (reduction to fuzziness). At this stage, the exact set of input data is converted into a fuzzy set, which is determined using membership functions.

2. Construction of the base of rules for fuzzy products.

3. Composition using aggregation methods.

4. Dephasing (reduction to clarity). At the stage of dephasing, the fuzzy system's executive module, on the basis of many fuzzy conclusions, forms an unambiguous decision with respect to the input variables.

Let us consider in more detail the initial stage on which phasing is carried out. Denote by d the input variable "suspended matter", which reflects the dustiness of the company's air environment. The corresponding term-set will be denoted by:

$$T1 = \{\text{low, medium, high}\} = \{D1, D2, D3\}.$$

The second input variable x - "chemical substances" - reflects the chemical contamination of the air environment. It corresponds to the analogous term set:

$$T2 = \{\text{low, medium, high}\} = \{X1, X2, X3\}.$$

The output variable y (the level of morbidity of workers) is also comparable to the analogous term set:

$$T3 = \{\text{low, medium, high}\} = \{Y1, Y2, Y3\}.$$

The next stage is the construction of a database of rules for fuzzy products. Most often, the Mamdani model is used as a model of fuzzy inference, the feature of which is the fact that its rules of inference contain fuzzy meanings in its consecutive clauses. In our case this is the membership function of the term-set $T3$.

In the chosen notation, we give for example some of these rules:

IF d IS $X1$ AND x IS $D1$ THEN y IS $Y1$

IF d IS $X2$ AND x IS $D1$ THEN y IS $Y2$

IF d IS $X3$ AND x IS $D1$ THEN y IS $Y3$

As a tool for implementing this approach, it is convenient to use the Fuzzy Logic Toolbox extension of the MATLAB computer mathematics environment, which allows creating fuzzy inference and fuzzy classification systems. The main interactive tool of the Fuzzy Logic Toolbox is the FIS inference editor, which contains tools for the functional mapping of input and output variables [8].

4. The discussion of the results

The stages of the simulation are illustrated in Figures 3-7 below.

Figure 3 shows a view of a software window with a schematic model.

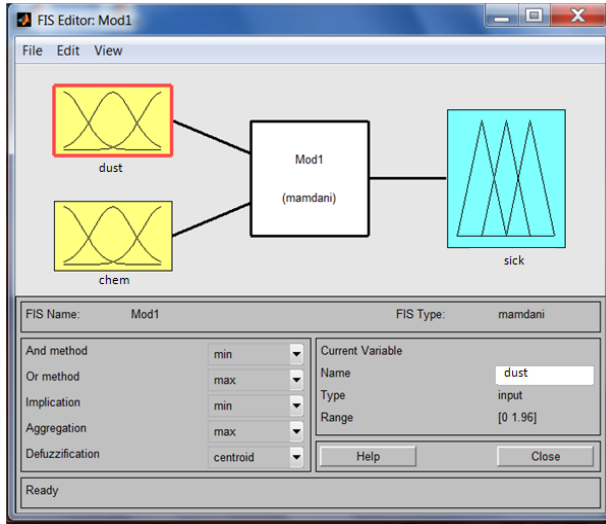


Fig. 3. Software window view

Figure 4 shows the graphs of the membership functions for the input terms "dust" (d), "chemical" (x) and output "sick" (y) of linguistic variables.

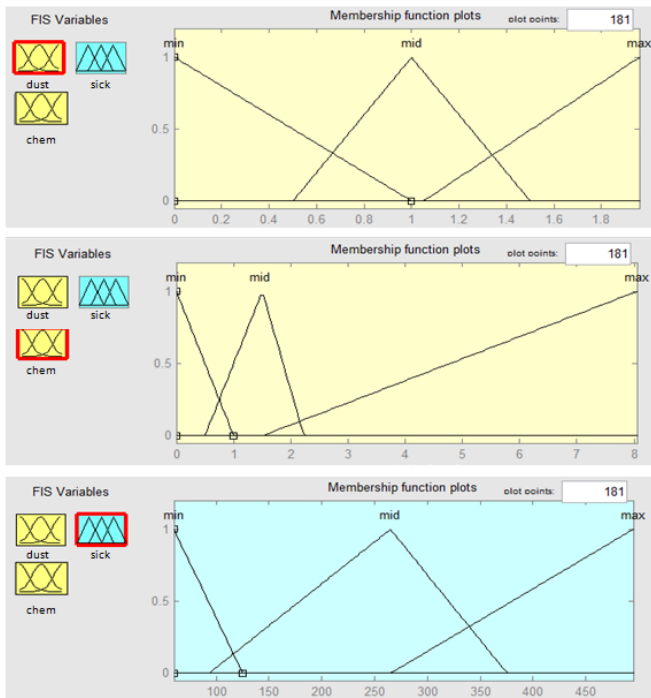


Fig. 4. Type of membership functions in the FIS editor of the Fuzzy Logic Toolbox extension package: "dust" and "chemical" - for indicators of non-carcinogenic risk for suspended and chemical substances, respectively; "sick" - for the number of cases of respiratory diseases per 1000 workers

Figure 5 shows a fragment of the window for determining the base of production rules for fuzzy inference.

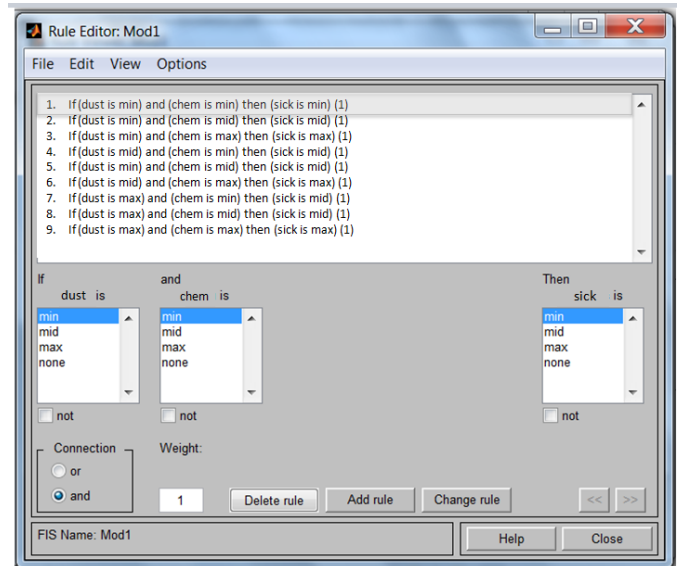


Fig. 5. The form of the product rules definition window in the FIS editor

Figure 6 shows the view of the fuzzy output viewer in the Mamdani model for the problem under consideration. Here, the aggregation of fuzzy rules is shown for two input variables, "sick" and "chemical". This uses a logical product, which corresponds to the operation min. Aggregation of implication concerning rules is carried out by logical summation, which corresponds to the operation max.

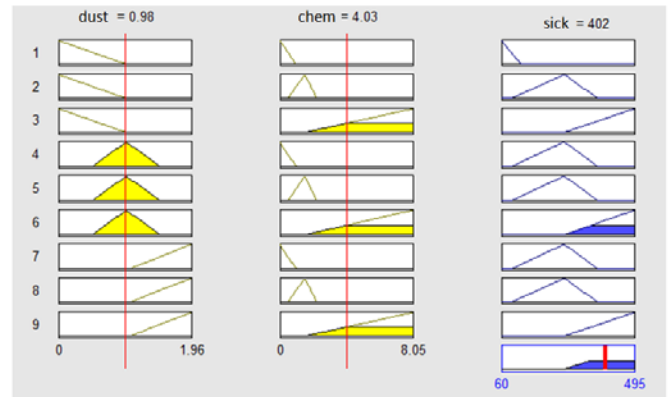


Fig.6. View of Mamdani's fuzzy output viewer

Since the max operator is used as an aggregation operator, and the min operator is used as an implication operator, the procedure for obtaining a fuzzy output value is a composition of max-min.

After receiving a fuzzy output (y), it is necessary to go to the phase of dephasing, which has the corresponding clear value y_{out} (4). As a method of dephasing, we used the method of the center of gravity:

$$(4) \quad y_{out} = \frac{\sum_{i=1}^n y_i \mu(y_i)}{\sum_{i=1}^n \mu(y_i)},$$

where $\mu(y_i)$ is the membership function of the i -th rule, and n is the number of fuzzy products rules.

Finally, Figure 7 shows the surface of the fuzzy output for the developed fuzzy model. This type serves for a general assessment of the adequacy of the constructed fuzzy model, and also allows analyzing the influence of the values of input variables on the value of the output variable.

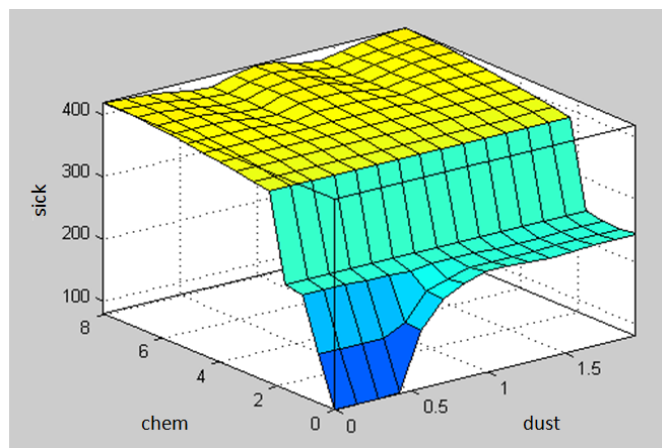


Fig. 7. The surface of fuzzy output for the developed fuzzy model

5. Conclusion

1. There are 29 professional groups of plant personnel, the data on which are converted into 2 input parameters d (suspended substances), x (chemical substances) and 1 output parameter y (the incidence of personnel with respiratory diseases).

2. A fuzzy model of the influence of pollutants of two groups on the health of personnel based on the Mamdani model, namely, on the respiratory system, has been developed and mathematically described.

3. A variant of phasing the three identified parameters based on an acceptable risk to staff health is suggested as "low", "medium", "high".

4. In the interactive mode, the fuzzy output system of the solved task is developed and visualized using the graphical tools of the Fuzzy Logic Toolbox extension package of the computer mathematics environment MATLAB.

5. It has been revealed that all personnel of the plant are exposed to a sufficiently large exposure to non-carcinogenic risk, the likelihood of harmful effects on the worker (respiratory diseases) increases in proportion to the increase in the coefficient of non-carcinogenic HI hazard.

6. The developed model can be easily supplemented by new indicators of air pollution or other production factors (linguistic variables) and new output parameters (fuzzy inference rules).

6. Literature

1. Assessment of non-carcinogenic risk for professional groups of "RN-Komsomolsk Oil Refinery" / I.V. Afanasyeva, V.V. Anisimov // Scientists notes Komsomolsk-on-Amur State Technical University. Sciences about nature and engineering, 2012. № II-1 (10). P. 89-96.

2. Identification of hazards in the assessment of health risks of personnel of RN-Komsomolsk Oil Refinery / I.V. Afanasyeva, V.V. Anisimov // Materials of the 11th International Scientific and Practical Conference in the Field of Ecology and Safety life activities "Far Eastern Spring - 2011". Komsomolsk-on-Amur : GOUVPO "KnAGTU", 2011. P. 141-149.

3. Diligenskiy N.V., Dymova L.G., Sevastyanov P.V. Fuzzy modeling and multicriteria optimization of production systems in conditions of uncertainty: technology, economics, ecology. M. : "Publishing Machine-Building-1", 2004. 398 p.

4. Mazorchuk M.S., Simonova K.A., Grekov L.D. Application of methods and models of fuzzy logic for modeling of economic processes // Systems of information retrieval, 2007. № 9 (67). P. 159-162.

5. Rogachev A.F., Melikhova E.V. Fuzzy-plural modeling and assessment of ecological safety of agricultural lands with radiation contamination // Global Nuclear Security, 2016. № 1 (18). P. 7-18.

6. Sereda S.N. Assessment of environmental risk through fuzzy models // Mechanical engineering and life safety, 2013. № 3. P. 15-20.

7. Tindova M.G. Fuzzy model of economic assessment of environmental damage // Economics: yesterday, today, tomorrow, 2012. № 3-4. P. 129-139.

8. Leonenkov A.V. Fuzzy modeling in MATLAB and fuzzyTECH. S-Pb. : BHV-Petersburg, 2005. 736 p.

SIMULATION MODELING OF AUDITORY FUNCTION

Ass. Prof., Dr. Eng. Donii O.¹, Dr. Med. Pisanko V.², Ass. Prof., Dr. Eng. Kulinich A.¹, Ass. Kotliar S.¹
National Technical University of Ukraine "Kyiv Polytechnic Institute named after Igor Sikorsky" - Kyiv, Ukraine¹
SI"O.S.Kolomyichenko Institute of Otolaryngology Of National Academy of Medical Science of Ukraine - Kyiv, Ukraine²
Email: dosha@iff.kpi.ua

Abstract: The hypotheses concerning encoding of information in the peripheral part of the human's auricular analyzer are presented. A brief critical analysis of contemporary trends in theoretical concepts concerning principles of work of the cochlea of the inner ear are conducted. Prerequisites for the construction of an alternative theory of coding of information in it are formulated in order to optimize the design and software of cochlear implants. The principle of constructing an imitation model of the generation of electric signals formed in the cochlea of the inner ear are proposed.

KEYWORDS: COCHLEAR, INNER EAR, BASILAR MEMBRANE, HEARING, SIMULATION MODELING

1. Introduction

Hearing plays an important role in the development of speech, intelligence and formation of the human's psyche. The information which is received through hearing, is not less important than the information which is perceived by sight. Loss of functions of hearing by a human largely limits his/her communication and leads to serious psychological and moral problems. Therefore, the task of full or partial, but socially adequate, recovery of hearing is an urgent one. To solve this problem hearing aids, as well as cochlear implants, are used in modern medicine. The task of hearing aids is simple and consists in shaping the frequency response of the amplifier to compensate for a decrease in the sensitivity of the patient's auditory system in a certain range of frequency. Implants of cochlea are constructed on the principle of separating the audio signal into several signals in separate frequency bands, followed by further direct electrical stimulation of the auditory nerve by each of them. They are used in case of violation of the peripheral part of the auditory analyzer, more specifically - of the patient's inner ear. Nowadays medicine has made significant progress in the use of cochlear implantation. Surgical techniques of their implantation are well established. However, there are certain difficulties related, mainly, with the perception of signals from the implant. These difficulties are caused by principles of implants' work, which are based on a fairly rough analysis of the audio signal's spectrum and selection of several frequency channels, each of which excites a certain portion of the auditory nerve. This ideology of the cochlear implantation is based on the current understanding of work of the peripheral part of the auditory analyzer, which is based on theory of Békésy [1], concerning the running wave on the basilar membrane (BM), and on "frequency-place" principle. There is enough experimental data that can be interpreted as a confirmation of this theory. However, there are also experimental data which contradict it [2]. One can say that today there is no consistent view on the functioning of the cochlea of the inner ear and representation of the audio signal in the structures, which stimulate the auditory nerve. Therefore, in this paper, a brief critical analysis of contemporary trends in the theoretical concepts of the peripheral part of the auditory analyzer is given and prerequisites are formulated for the construction of an alternative hypothesis on the coding of information in auditory analyzer. Creation of the experimentally proved theory on the basis of this hypothesis will provide an opportunity for further critical rethinking of principles of the cochlear implants' design in order to improve the adaptation of patients with deafness caused by damage of the inner ear's cochlea.

2. Preconditions for resolving the problem

One of the known effects, inexplicable from the standpoint of the theory of "frequency-place" is the binaural effect. Evaluation of the time difference of arrival of the same wave's phases to both ears may occur, evidently, only in the brain centers, that means that the periodic nature of the sound process should somehow appear in the neural processes of the cortex. Meanwhile, the theory of "frequency-place", as the theory of "peripheral analyzer" refers assessment of sound solely to the excitation of nerves in the given area of the cochlea. This leads to the emergence of new theories of hearing. One of such theories is the theory of G. Fletcher [3].

According to this theory, it is not individual strings of basic membrane respond to the audio waves, but peri- and endolymph of the cochlea do this. Plate of the stapes transmits sound vibrations of the cochlea's fluid to the BM, at that the maximum of amplitude of these oscillations at higher tones lies closer to the base of the cochlea, at lower ones - closer to its top. Nerve fibers which end in the main membrane, including the organ of Corti (according to some authors) resonate only at frequencies above 60 - 80 Hz. There are no fibers, which receive more lower frequency on the main membrane. Nevertheless, feeling of heights up to 20 Hz is formed in the brain. It appears like the combination of high tone harmonies. Thus, from the Fletcher hypothesis's viewpoint, perception of the low tones' pitch is explained by perception of the whole complex of harmonic overtones, and not only by perception of the frequency of the main tone, as it was usually taken so far. And as content of overtones to a large extent is dependent on the intensity of the sound, then it becomes clear a close relationship between the three subjective qualities of the sound - its height, volume and timbre. All these elements, each of them individually, are dependent on the frequency, strength and composition of the sound's overtones. According to the Fletcher's hypothesis, resonant properties are inherent to the mechanical system of the cochlea as a whole, not just to the main membrane's fibers. Under the influence of a certain pitch, not only fibers, which resonate with this frequency, oscillate, but the entire membrane as well, and also this or that amount of fluids in cochlea. High tones force to drive only a small mass of liquid near the base of the cochlea, low ones - are fixed closer to helicotrema. Fletcher also overcomes the main difficulty of the resonance theory associated with explanation of a large range of volume. He believes that the volume is determined by the total number of nerve impulses, coming to the brain from all the excited nerve fibers of the basal membrane. Fletcher's theory, in general, does not deny existence of "frequency-place" theory and it can be attributed to the theories of "peripheral analyzer".

Theories of "central analyzer" or so-called "telephone theory", form another group of theories [4]. According to these theories, audio vibrations are converted by cochlea into synchronous waves in the nerve and transmitted to the brain, where their analysis and perception of level of tone takes place. J. Ewald's theory, which was proposed in the late of 19th century, also belongs to this group of theories. According to this theory, under effect of the sound, standing waves are formed in the cochlea with a length which is determined by the frequency of the sound. The level of the tone is determined by the perception of the shape of the pattern of the standing waves. The feeling of a certain tone corresponds to the excitation of one part of the nerve fibers, and the feeling of a different tone corresponds to the excitation of another part. Analysis of sound is performed not in the cochlea but in the central areas of the cortex. Ewald succeeded to build a model of BM, with the size which approximately corresponded to the real ones. In his experiments the entire membrane began to vibrate when it was excited by the sound. The "audio picture" appears in the form of standing waves with the length, which is as smaller, as the sound is higher. Despite the successful explanation of some embarrassing particulars, Ewald's theory (as well as other theories of "the central analyzer") hardly corresponds to the latest physiological researches

of the nerve impulses' nature. In the [3] the dual point of view is expressed, namely, the explanation of the perception of high tones is given in the sense of "peripheral analyzer" and of low ones - from the perspective of "the central analyzer".

American researchers, who first-ever implanted microelectrodes in cat's cochlea, registered electrical potentials which arised in the cochlea [5]. Based on their observations, they created an electrophysiological theory of hearing. According to this theory, every hair of hair cells of organ of Corti is similar to the piezoelectric crystal. As it is known, these crystals have an interesting property - upright they are neutral, but when they are bent even a little bit then electric charge appears immediately. In case of fluctuations of BM, hair cells, naturally, begin to oscillate also. But the tectorial membrane pushes on top of the hair, they bend, causing an electrical charge. Thus, under the influence of the deformation of receptor cells' hairs in sync with the sound's vibrations, the electrical energy is released, and biological currents appear. These biological currents stimulate the thinnest endings of branches of auditory nerve, which criss-crosses the hair cells. Through this nerve and conductive pathways of medulla oblongata excitation is transferred to the cortex of the temporal lobes of the brain, where the analysis and synthesis of audio stimuli takes place. Thus, at the moment it is common to speak about the duality of the mechanisms of perception of the pitch: in the high-frequency range the most acceptable is the principle of "place", in the area of lower frequencies - a modified principle of "bursts". Despite the long history of the discussions on the principles of the functioning of the auditory system as a whole and its peripheral parts in particular, and the availability of a huge number of researches related to the study of the perception of pitch, it is obvious that the mechanism of coding the information in the peripheral part of the auditory analyzer is not completely elucidated and requires further intensive research.

3. Formulation of a hypothesis for model development

From the perspective of physics, dynamic range of 125 dB is a unique parameter of the human auditory system. Thus, the maximum amplitude of the audio signal at the system's input (eardrum) differs from the minimal amplitude in trillion times. Such value of the dynamic's range leads to suggestion that there is not only one, but several mechanisms of perception of the sound in the peripheral part of the auditory analyzer, which is consistent with the views of a number of scientists. For example, it is known [6], that the subcutaneous plate of stirrup at high intensity of the input's signal, close to the maximum intensity, moves from the translational vibrations to the vibrational-rotational ones, thus preventing the entire system of middle and inner ear from the mechanical damage. Then this is logical to assume the existence of several reactions to sounds, which are different not only by frequency, but also by intensity. Thus, it is interesting to compare processes of formation of auditory images in the cochlea of the inner ear at maximum and minimum levels of input signals.

Taking into account the assumption that there are several mechanisms of the inner ear's functioning while converting sounds of different intensity into the electrical activity of the auditory nerve, let us consider the mechanical processes in the cochlea when it is exposed to weak (low intensity) beeps.

As it was mentioned above, Békésy G. substantiated hypothesis, called the "theory of running waves", which states that the vibrations are distributed through BM in the form of running waves having a maximum of the envelope at a certain point of membrane, whose place varies depending on frequency of the acting signal. A number of authors who used different methods for registration of membrane's oscillations, experimentally confirmed the basic provisions of G. Békésy's hypothesis [7 - 10]. However, direct observations of BM's vibrations, executed by G. Békésy, were carried out at a power of the sound of 90 - 120 dB. Using an extrapolation of the existing experimental data it is possible to calculate that the amplitude of BM's movement in the place of the running wave's envelope's maximum is about 10^{-12} sm [1, 7]. This value is much smaller than the amplitude of the thermal motion of

the molecules of cochlea fluids. Furthermore, in order to have any fluctuations of BM, the pressure of sound's signal must overcome cochlea's elasticity and its inertia of rest. And, obviously, because of its stiffness, it will not respond to such a small amount of energy. So, it can be postulated that BM is stationary at low intensities of the sound's signal. And this gives a basis to suppose the absence of influence of BM's mechanics on the analysis of the sound's information at small acoustic signals. At the same time, the presence of a running wave can say about ability of the cochlea's structures to absorb the excess of energy at high levels of the sound. This assumption, as well as high sensitivity of acoustic analyzer, give reasons to look for other mechanisms of perception of the sound in the cochlea, especially at low levels of intensity of acoustic signals. In a number of works [8, 9], where the amplitude-frequency characteristics of the BM's vibrations were investigated, it turned out that they are less selective than it was expected based on the values of differential threshold of frequency. This allowed to suppose the presence of so-called "the second filter", which is located, according to the researchers' viewpoint, at the junction of the tectorial membrane - wax cells of the organ of Corti or hair cells - the fibers of the auditory nerve Ψ [11 - 13], although such a filter was not detected experimentally [11]. Also in the work [14] the hypothesis is suggested concerning the existence of molecular resonance mechanism, which is localized in the tectorial membrane. Due to different speeds of sound's movement in the perilymph and in the tectorial membrane, concentration of the acoustic energy takes place in the latter, resulting in the conversion of mechanical energy into biochemical one through resonance and movement of ions, which cover complex of the tectorial membrane - the hair cells of organ of Corti. There is a hypothesis, according to which the flanking sprouts of Deiters cells and cuticular region of the hair cell form a pair, called "Auron" by the authors, that makes primary frequency analysis of sounds using resonant oscillations (hypothesis of "auron resonance"). A number of authors [4, 15] express the opinion that the frequency analysis in the cochlea is carried out by means of resonant vibrations of the hair cells and the tectorial membrane. These assumptions are confirmed by correlation between the lengths of hair cells' stereocilia and characteristic frequencies [4, 15]. Thus, modern researches have led to the need for a thorough study of the "thin" structures of the organ of Corti.

Imagine the possible mechanism of the cochlea's mechanical part under the influence of weak acoustic signals (0-20 dB above threshold). In this case, taking into account the above mentioned extrapolation, let us assume that BM is immovable. If we are adhere to the concept of a mechanical nature of transformation of sound's vibrations into receptors' potentials of the hair cells, it is necessary to find in the structure of the organ of Corti those oscillating elements, that have to "feel" the impact of such a small incentive. To do this, we consider a schematic cross-sectional view of the organ of Corti in a single turn of the cochlea (Figure 1). The figure shows that the hair cells are connected with the tectorial membrane (this is confirmed experimentally by [1]), and the latter, in turn, is fixed in such a way that it can form a very sensitive lever system in vertical direction regarding to the BM's plane. The schematic connection of the covering membrane with hair cells is shown in Figure 2, at that it is assumed that the hair cells have their own elasticity (coefficients K_j). When the mechanical structure of the cochlea's membrane was studied, it was found that the tectorial membrane really moved easily in the direction which is perpendicular regarding to BM [1].

In the experiments described in the works of [16], it is shown that the stereocilia of hair cells are rather hard. It can be assumed that they alone or in a bundle together with the mass of the coating membrane form sensitive, and perhaps the resonance system. In the [4, 15], for example, the possibility of resonance in such a system is shown. This is confirmed also by the structure of the coating membrane: it consists of thin transverse fibers, which generally have a radial direction along the axis of the cochlea, and there is a transparent gluing substance between the the fibers [17]. Reissner's membrane is a very thin film of the same elasticity over the entire length. At that, its elasticity is small in comparison with the

elasticity of BM, through which transmission of oscillations passes freely. Its role is to separate endo- and perilymph and it seems, that it does not participate in the analysis of vibrations.

Based on the above mentioned facts, one can imagine a process of the membranes' vibrations in the cochlea in terms of action of acoustic signals with low levels of sound's pressure as follows: sound's pressure is transferred to a tectorial membrane through the perilymph, endolymph and Reissner membrane. Because of its structure, tectorial membrane responds to minimum pressure by substantial displacement in a plane, which is perpendicular to the BM, as compared with the displacement of MB (or its absence) at the same levels of input signals. The mass of the tectorial membrane and elasticity of the hair cells can form the sensitive system, which reacts to this movement. Moreover, taking into account the structure of the tectorial membrane, it can be assumed, that together with the hair cells, it forms the whole system of sensitive elements, which, in principle, can work as unbound tuned resonators. Taking into account calculations contained in [4, 15], their amplitude and frequency characteristics will be sharper than the same specifications for MB. BM in this case remains stationary and does not participate in the analysis of signals with such levels. It is obvious, that while increasing the sound's level, starting from some particular magnitude of the acoustic signal, vibrations of the BM become relevant and can not be ignored anymore. In this case the nature of excitation of hair cells, which determines the information transmitted by the auditory nerve, is to be altered.

This situation can be considered in more details. Under the influence of a weak signal BM is stationary. Sound's pressure, according to the laws of physics, is transmitted through the perilymph in all directions. At that, if the forward movement of the stapes is very slow (very low frequency), a column of liquid should overflow from channel to channel through helicotrema without creating significant pressure on the side walls. In this case, hair cells do not respond to the signal. By increasing the speed of movement of the stapes (increasing the frequency of the signal) perilymph does not have time to go through an opening with a limited area (helicotrema) and the radial component of pressure arises, which affects the complex "tectorial membrane - hair cells", generating an electric signal (it is possible that the area of helicotrema defines the lower boundary of perception of the signal). This pressure should be distributed over the entire length of the channel, that corresponds to the assumptions of the "telephone theory"

Thus, when the input signal has low intensity, the dotted excitation response of hair cells is possible. However, considering the fact, that pressure in liquid circulates in all directions simultaneously, and that the helicotrema restricts the flow of the perilymph from channel to channel, one may assume the simultaneous stimulation of hair cells in the area of a certain length.

The increase in signal intensity leads to the excitation of BM's vibrations and the appearance of a running wave on it. A running wave should lead to the displacement of structures of organ of Corti and irritation of hair cells throughout its whole length, not just at the point of its maximum amplitude. So, it is possible to propose a hypothesis about the formation of a some spatio-temporal pattern, which represents the excitation of the auditory nerve. At that, taking into account the form of envelope of a running wave, the excitation should reflect the reaction of hair cells in three spatial coordinates, as well as changes in the spatial "images" in time. Thus the spatial-temporal signal, having four coordinates: length, width, depth and time, is formed. Obviously, in such way principle of "frequency-place" and elements of the telephone theory are combined.

Considering the mechanisms of forming the excitation in the cochlea of the auditory analyzer, the question must arise about the nature of the signal applied to the input of the system in experimental studies. Usually it is a pure tone, which represents a sine wave with a certain amplitude and frequency. It is believed that the "frequency-place" principle works in this case. When the composite signal is supplied, then since the Helmholtz times in most studies it is suggested that the ear analyzes in this or that way its spectrum, represented as a set of sinusoids. However, representation of the composite signal as a set of sine and cosine,

which is widely used in electronics and acoustics, is a comfortable way, but not the only one. For example, according to the approximation theorem of Weierstrass any complex function can be approximately described by a polynomial of degree n , i.e. as a sum of exponential functions with different coefficients of [18]. And taking into account the fact that in the natural environment you can hardly meet pure tones, the natural response of the auditory system is the analysis of complex sounds. It is doubtful the possibility of the membrane of the cochlea of the inner ear to decompose non-periodic audio signal in Fourier series. Thus, it is not necessarily that the inner ear works as a spectrum analyzer, allocating and fixing the harmonic components of the audio signal. Hence the assumption arises (which supports the expressed above hypothesis) that the audio signal is not decomposed in the inner ear into components, but is perceived as a whole. Of course, the principle of "frequency-place" can not be denied, as it was confirmed experimentally. However, its confirmation was received when exposed to high intensity signals.

4. Computer model of signal formation in the cochlea

These assumptions should be confirmed by experimental verification. At the same time, it should be noted that the experimental study of structures (and particularly of "thin" structures) of the inner ear's cochlea in real objects are very labour-consuming and not very informative, even when using modern technology and equipment. Therefore simulation is one of the main methods of investigation of the inner ear. Unfortunately, among a large number and variety of existing models of cochlea, there are just a few of them which take into account not only the behavior of the BM, but also of other membranes, but even at that case they are used only with the aim to clarify its vibrational characteristics.

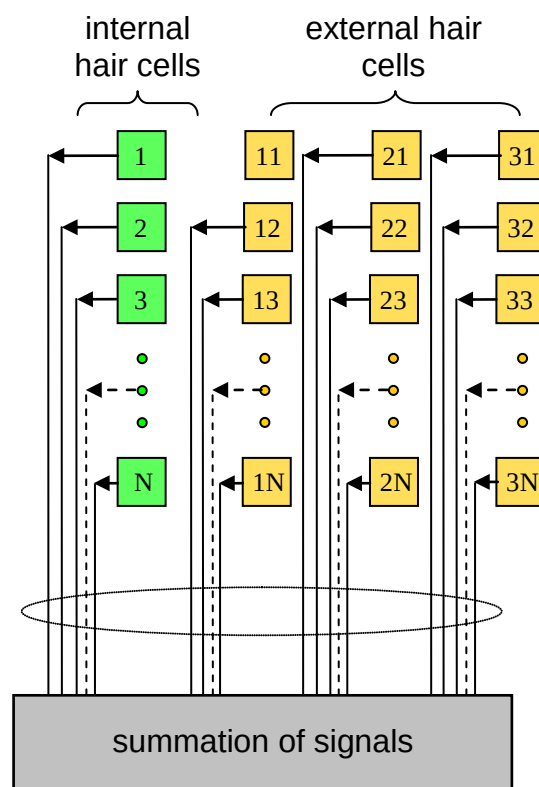


Fig. 1. Structure of the imitation model of the inner ear function

Taking into account complexity of the structure of the inner ear's cochlea, simulation modeling is conveniently chosen as a method of constructing of the model, since this technique was developed for studying of complex systems. Organization of hydromechanical part of the cochlea is complex, not linear and poorly researched. However, its response to the sound signal causes the excitation of hair cells, both internal and external. This excitation forms a common electrical signal, which is transmitted to the upper parts of

the brain along the auditory nerve. Based on the proposed hypothesis of the formation of some spatio-temporal pattern of oscillations in the cochlea, it is possible to model different variants of hair cells' excitation, and, accordingly, of their reaction as an aggregated electrical signal. This electrical signal can be registered in a real experiment. Comparing signals which were simulated by the model with the experimentally recorded ones, one can select the most similar ones and thereby make a conclusion about the real mechanisms of the inner ear's functioning. A simplified diagram of the model of the formation of an electrical signal of the hair cells' reaction to an external signal is shown in Fig. 1. Inner and outer hair cells are denoted by multicolored squares. In this variant, each cell (which can be regarded as a sensor) generates a single pulse regardless of the parameters of the input signal. All pulses which were generated at a given moment of time are summed in the summation block. Thus, a signal is generated, which is similar to that one, which can be registered in a real experiment. The fragment of output results of modeling of hair cells' excitation under influence of the input signal which was generated by means

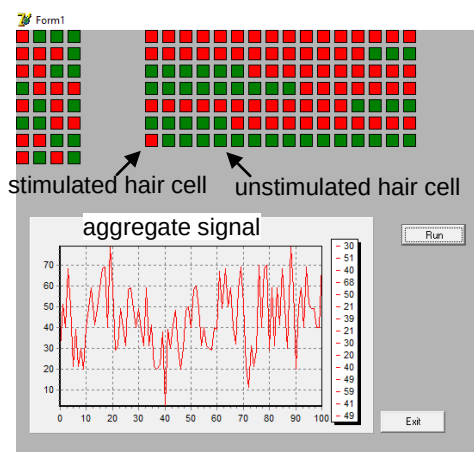


Fig. 2. The computer simulation result of the formation of an electrical signal at the "exit" of the cochlea of the inner ear

of the generator of random numbers is presented in Fig. 2. As it can be seen from Fig. 2, the resulting signal is similar to the microphone potential which was measured in a real experiment. Thus, one can see the prospects of developing of this approach for modeling of the auditory function. One can draw a conclusion about the principles of coding of information in the peripheral part of the human auditory system by analyzing the mechanisms of functioning of the inner ear's cochlea from the point of view of existing theories and developing models of signals that stimulate hair cells and comparing the results of modeling with the results of the experiment. Moreover, it seems expedient to select tests of subjective and objective researches of hearing, on the basis of which it is possible to create an adequate model of inner ear's functioning both in conditions of norm and pathology. This will deepen our knowledge of coding of auditory information and optimize design and software of cochlear implants on their basis. In case of confirmation of mechanisms of the inner ear's functioning in framework of the above mentioned hypothesis, the design of cochlear implants must be based on other ideas about coding of signals in the peripheral part of the auditory analyzer other than decomposition of the input signal into several frequency ranges.

5. Conclusion

1. The hypothesis of stimulation of hair cells of the inner ear's cochlea is proposed, in which the formation of a spatio-temporal picture of excitation as a reaction to an external sound signal is suggested. This hypothesis makes it possible to explain the sensitivity of the auditory analyzer at low intensities of input signal and, on the whole, does not contradict to the existing ideas about mechanisms of hearing's perception.
2. An imitation model of formation of an electrical signal of the hair cells' reaction to an external signal is proposed, in which each

cell generates a pulse and all pulses are summed in the summation block. In this way the electric reaction of a cochlea is simulated, which is similar to that one, which can be recorded in a real experiment.

3. It seems appropriate to conduct tests of subjective and objective researches of hearing for the analysis of modeling results on the basis of which the adequacy of the proposed hypothesis of functioning of the inner ear can be confirmed in conditions of norm and pathology. This will allow to deep knowledge about coding of auditory information and optimize on this basis design and software of cochlear implants.

6. Literature

1. Bekesy G. Experiments in Hearing. N.Y., McGraw-Hill, 1960.
2. Rubinshtein S.L. Foundations of General Psychology, St. Petersburg, Peter, 2002. (Russian)
3. Fletcher H. On the dynamics of the cochlea. - J. Acoust. Soc. Amer., 1951, v. 23., p. 637 - 645.
4. Zwislocki, J.J. Theory of cochlear mechanics. Hear. Res., 2, 1980, P. 171 - 182.
5. Kiang N.Y-S., Watanabe T., Thomas E.C., Clark L.F. Discharge patterns of the single fibers in the cat's auditory nerve. Cambridge, Massachusetts. 1965.
6. Gelfand S.A. Hearing: An Introduction to Psychological and Physiological Acoustics, Marsel Dekker, INC, New York and Basel, 1981.
7. Johnstone B.M., Taylor K.J., Boyle A.J. Mechanics of the guinea-pig cochlea. J. Acoust. Soc. Amer., 47, 1970, p. 504 - 509.
8. Johnstone B.M., Yates G.K. Basilar membrane tuning curves in the guinea pig. J. Acoust. Soc. Amer., 55, 1974, p. 584 - 587.
9. Kohlöffel L.U.E. Observations of the mechanical disturbances along the basilar membrane with laser illumination. In: Basic mechanisms in hearing. New York; London, 1973, p. 95 - 117.
10. Rhode W.S. Some observations on cochlear mechanics. J. Acoust. Soc. Amer., 64, 1978, p. 158 - 176.
11. Labutin V.K., Molchanov A.P. Models of the mechanisms of hearing. M., 1973. (Russian)
12. Duifhuis D. Cochlear linearity and second filter: Possible mechanisms and applications. J. Acoust. Soc. Amer., 59, 1976, p. 356 - 358.
13. Hall J.L. Spatial differentiation as an auditory "second filter": assesment on a nonlinear model of the basilar membrane. J. Acoust. Soc. Amer., 61, 1977, p. 520 - 524.
14. Naftalin L., Path F.R.C. The Peripheral Hearing Mechanism: A Biochemical and Biological Approach. Annals of Otolaryngology, Rhinology & Laryngology, v. 85, №1, 1976, p. 38 - 42.
15. Zwislocki, J.J. Five decades of research on cochlear mechanics. J. Acoust. Soc. Amer., 67, 1980, p. 1679 - 1685.
16. Flock A. Physiological properties of sensory hairs in the ear/ In: Psychophysics and physiology of hearing. London, 1977.
17. Vinnikov Y.A., Titova L.K. Corti's organ. Histophysiology and histochemistry. M.-L., 1961. (Russian)
18. Fikhtengolts G.M. Course of differential and integral calculus. FIZMATLIT, v. 3, 2001. (Russian)