

НАУКОВІ ВІСТІ КПІ

Міжнародний науково-технічний журнал

№ 4 (134)

2021

Започаткований у вересні 1997 року

Головний редактор
М. З. Згуровський, акад. НАН України

Заступник головного редактора
М. Ю. Ільченко, акад. НАН України

Відповідальний секретар
П. П. Маслянюк, канд. техн. наук, доц.

У номері:

Інформаційні технології,
системний аналіз та
керування

Матеріалознавство та
машинобудування

Приладобудування та
інформаційно-вимірювальна
техніка

Адреса редакції:
КПІ ім. Ігоря Сікорського
просп. Перемоги, 37,
Київ, 03056, Україна
Тел. (+38 044) 204-91-23
E-mail: n.visti@kpi.ua
<http://scinews.kpi.ua>

Засновник – Національний технічний університет України
“Київський політехнічний інститут імені Ігоря Сікорського”

Свідоцтво про державну реєстрацію журналу – серія КВ № 23144-12984ПР, видане 7 серпня 2017 року

Згідно з наказами МОН України № 1643 від 28.12.2019 та № 409 від 17.03.2020 журнал включено до категорії “Б” Переліку наукових фахових видань України з технічних наук (спеціальності — 113 Прикладна математика, 121 Інженерія програмного забезпечення, 122 Комп’ютерні науки та інформаційні технології, 123 Комп’ютерна інженерія, 124 Системний аналіз, 131 Прикладна механіка, 132 Матеріалознавство, 133 Галузеве машинобудування, 134 Авіаційна та ракетно-космічна техніка, 141 Електроенергетика, електротехніка та електромеханіка, 142 Енергетичне машинобудування, 143 Атомна енергетика, 144 Теплоенергетика, 151 Автоматизація та комп’ютерно-інтегровані технології, 161 Хімічні технології та інженерія, 171 Електроніка, 172 Телекомунікації та радіотехніка)

Рекомендовано Вченою радою Національного технічного університету України
“Київський політехнічний інститут імені Ігоря Сікорського”, протокол № 3 від 14.02.2022 р.

Члени редакційної колегії

М. І. Бобир,	д-р техн. наук, проф., чл.-кор. НАН України
Є. Бондарев,	PhD, проф., Нідерланди
Х. Валеро,	PhD, проф., Іспанія
І. А. Дичка,	д-р техн. наук, проф., Україна
П. І. Лобода,	д-р техн. наук, проф., акад. НАН України
Г. С. Тимчик,	д-р техн. наук, проф., Україна
П. Хенаф,	д-р, проф., Франція
О. Е. Чигиринець,	д-р техн. наук, проф., Україна

Секретар редакції **Г. В. Воронка**

Це останній номер журналу нашої колеги Ганни Воронки.
24 лютого 2022, Ірпінь, Україна

Редактор В. М. Ігнатович, **Г. В. Воронка**

Комп’ютерна верстка С. А. Бобров

Національний технічний університет України
“Київський політехнічний інститут імені Ігоря Сікорського”
Свідоцтво про державну реєстрацію: серія ДК № 5354 від 25.05.2017 р., просп. Перемоги, 37, Київ, 03056

Підп. до друку 30.08.2022. Формат 60×84¹/₈. Папір офс. Гарнітура UkrainianTimesET.
Спосіб друку – електрографічний. Ум. друк. арк. 12,55. Обл.-вид. арк. 17,42. Наклад 50 пр. Зам. № 00-000.

Видавництво “Політехніка” КПІ ім. Ігоря Сікорського
вул. Політехнічна, 14, корп. 15, Київ, 03056
тел. (044) 204-81-78

KPI SCIENCE NEWS

International research journal

№ 4 (134)

2021

Founded in September, 1997

Editor-in-chief

M. Z. Zgurovsky, Academician of NASU

Deputy editor-in-chief

M. Yu. Ilchenko, Academician of NASU

Executive editor

P. P. Maslyanko, Assoc. Prof., PhD

In the issue:

Information technologies, system
analysis and control

Science of materials and
mechanical engineering

Instrument-making and
information-measuring equipment

Editorial office:

Igor Sikorsky Kyiv Polytechnic Institute,
building 1, Ave. Peremogy, 37,
Kyiv, 03056, Ukraine

E-mail: n.visti@kpi.ua

<http://scinews.kpi.ua>

Founder – National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”
Registration Certificate – KB No 23144-12984ПП

*According to the order of MES of Ukraine from 28.12.2019 no. 1643 is a specialized journal on technical sciences
(specialities – 113, 121, 122, 123, 124, 141, 142, 143, 144, 151, 161, 171, 172)*

Advised by the Academic Council of the National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”, protocol No 3 on 14.02.2022.

Editorial Board

Mykola Bobyr,	Prof., Correspondent member of NASU, Ukraine
Egor Bondarau,	Prof., Netherlands
Olena Chyhyrynets,	Prof., Ukraine
Ivan Dychka,	Prof., Ukraine
Patrick Henaff,	Prof., France
Petro Loboda,	Prof., Academician of NASU, Ukraine
Grygorii Tymchik,	Prof., Ukraine
José Valero,	Prof., Spain

Editorial secretary A. V. Voronka

This is the last issue of our colleague Hanna Voronka’s magazine.
February 24, 2022, Irpin, Ukraine

Editor V. M. Ihnatovycn, A.V. Voronka

Desktop publishing S. A. Bobrov

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”
Registration Certificate – ДК № 5354 on 25.05.2017, Ave. Peremogy, 37, Kyiv, 03056

Signed for printing on 30.08.2022. Format 60×84¹/₈. Text-weight paper. Font UkrainianTimesET.
Print. tech. – electrographic. Convent. printed sheets 12,55. Published sheets 17,42. Edition of 50 copies. Order No 00-000.

Publishing House “Politehnika”, Igor Sikorsky Kyiv Polytechnic Institute
Politekhnichna Str., 14, building 15, Kyiv, 03056
Tel.: (044) 204-81-78

З М І С Т

Інформаційні технології, системний аналіз та керування

<i>Kuznetsova N.V., Trofymchuk O.M., Bidyuk P.I., Terentiev O.M., Levenchuk L.B.</i> Bayesian modeling of risks of various origin	7
<i>Romanuke V.V.</i> Finite approximation of zero-sum games played in staircase-function continuous spaces....	19
<i>Маслянюк П.П., Савчук І.В.</i> DevOps – концепт і структурне представлення	39

Матеріалознавство та машинобудування

<i>Кузнєцов Ю.М., Сінмінь Г., Самійленко О.В.</i> Передумови використання системно-морфологічного підходу та теорії фракталів при створенні лещат для об'єктів складної форми	52
<i>Ремізов Д.О., Степанчук А.М., Соловійова Т.О., Лобода П.І., Завадюк С.В.</i> Використання самофлюсівних сплавів як джерела легувальних елементів під час отримання порошкових виробів конструкційного призначення з низьколегованого заліза	58
<i>Середа В.В., Соломаха А.С., Притула Н.О., Заболотний О.А.</i> Термодинамічний аналіз термічної опріснювальної установки з циклом зволоження–осушення повітря.....	69
<i>Трокоз Я.Є., Барабаш П.О., Голубєв О.Б., Куделя П.П.</i> Розроблення та дослідження холодильного циклу з біагентним струминним компресором.....	77
<i>Янчевський І.В., Нехороших М.В.</i> Вплив геометричних параметрів кривошипно-планетарного редуктора на циклограму роботи контактних пар	83
<i>Мацевитий Ю.М., Бушетець Я.М.</i> Різновиди та технічні характеристики сонячних колекторів	92

Приладобудування та інформаційно-вимірювальна техніка

<i>Grigoriy S. Tymchik, Oleksandr O. Podolian, Maksym S. Druzhev</i> Controlling the thickness of the plastic layer applied using 3D printing by electromagnetic acoustic method	99
Автори номера.....	106

CONTENTS

Information technologies, system analysis and control

<i>Kuznietsova N.V., Trofymchuk O.M., Bidyuk P.I., Terentiev O.M., Levenchuk L.B.</i> Bayesian modeling of risks of various origin	7
<i>Romanuke V.V.</i> Finite approximation of zero-sum games played in staircase-function continuous spaces....	19
<i>Maslianko P.P., Savchuk I.V.</i> DevOps – concept and structural representation	39

Science of materials and mechanical engineering

<i>Kuznetsov Yu. M., Xingmin G., Samoilenko O. V.</i> Prerequisites for using the system-morphological approach and fractal theory in the creation of vices for complex shape objects	52
<i>Remizov D.O., Stepanchuk A.M., Soloviova T.O., Loboda P.I., Zavadiuk S.V.</i> The use of self-fluxing alloys as a source of alloying elements in producing powder products of construction purpose from low-alloyed iron	58
<i>Sereda V.V., Solomakha A.S., Prytula N.O., Zabolotny O.A.</i> Thermodynamic analysis of thermal desalination system with humidification–dehumidification cycle	69
<i>Trokoz Ya.Ye., Barabash P.O., Golubev O.B., Kudelya P.P.</i> Development and research of refrigeration cycle with biagent jet compressor	77
<i>Yanchevskiy I.V., Niekhoroshykh M.V.</i> Influence of geometric parameters of crankshaft-planetary reducer on cyclograph of work of contact pairs	83
<i>Matsevytyi Yu.M., Bushtets Ya.M.</i> Varieties and technical characteristics of solar collectors	92

Instrument-making and information-measuring equipment

<i>Grigoriy S. Tymchik, Oleksandr O. Podolian, Maksym S. Druzhev</i> Controlling the thickness of the plastic layer applied using 3D printing by electromagnetic acoustic method	99
Contributors to the issue	106

DOI: 10.20535/kpissn.2021.4.251684

UDC 004.942:519.216.3

N.V. Kuznietsova¹, O.M. Trofymchuk², P.I. Bidyuk^{1*}, O.M. Terentiev², L.B. Levenchuk¹¹ Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine² Institute of Telecommunications and Global Information Sphere at NAS of Ukraine

*corresponding author: pbidyuke_00@ukr.net

BAYESIAN MODELLING OF RISKS OF VARIOUS ORIGIN

Background. Financial as well as many other types of risks are inherent to all types of human activities. The problem is to construct adequate mathematical description for the formal representation of risks selected and to use it for possible loss estimation and forecasting. The loss estimation can be based upon processing available data and relevant expert estimates characterizing history and current state of the processes considered. An appropriate instrumentation for modelling and estimating risks of possible losses provides probabilistic approach including Bayesian techniques known today as Bayesian programming methodology.

Objective. The purpose of the paper is to perform overview of some Bayesian data processing methods providing a possibility for constructing models of financial risks selected. To use statistical data to develop a new model of Bayesian type so that to describe formally operational risk that can occur in the information processing procedures.

Methods. The methods used for data processing and model constructing refer to Bayesian programming methodology. Also Bayes theorem was directly applied to operational risk assessment in its formulation for discrete events and discrete parameters.

Results. The proposed approach to modelling was applied to building a model of operational risk associated with incorrect information processing. To construct and apply the model to risk estimation the risk problem was analysed, appropriate variables were selected, and prior conditional probabilities were estimated. Functioning of the models constructed was demonstrated with illustrative examples.

Conclusions. Modelling and estimating financial and other type of risks is important practical problem that can be solved using the methodology of Bayesian programming providing the possibility for identification and taking into consideration uncertainties of data and expert estimates. The risk model constructed with the methodology proposed illustrates the possibilities of applying the Bayesian methods to solving the risk estimation problems.

Keywords: financial processes; financial risks; Bayesian programming methodology; risk estimation.

Introduction

Repeating financial crises, unfavourable changes of climate, local military conflicts between many countries of the world and fighting terrorists in multiple locations give an evidence for high necessity of solving the problems of analysis and management of growing various type risks in every area of human activity. Together with these highly unfavourable for economic and social developments events the facts are revealed that existing methods of risk analysis, and modelling the situations concentrated on creating their formal description and appropriate

risk management procedures usually come with some delays or are inadequate for the quality risk estimating and forecasting in conditions of multiple random external disturbances (risk factors). This is mostly explained by the high dynamics of modern processes (especially financial ones, ecological and climate changes), their high dimensionality, sophisticated vertical and horizontal interconnections at the level of separate companies, economy branches, macro-economy as a whole and at the global level as well [1]. Actually most of the risks met in everyday life, including the risks of natural disasters, can be analysed from the financial point of view, and this

Рекомендуємо цитувати цю статтю так: N.V. Kuznietsova, O.M. Trofymchuk, P.I. Bidyuk, O.M. Terentiev, L.B. Levenchuk, "Bayesian modelling of risks of various origin", *Наукові вісті КНІ*, no. 4, pp. 7–18, 2021. doi: 10.20535/kpissn.2021.4.251684.

Please cite this article as: N.V. Kuznietsova, O.M. Trofymchuk, P.I. Bidyuk, O.M. Terentiev, and L.B. Levenchuk "Bayesian modelling of risks of various origin", *KPI Science News*, no. 4, pp. 7–18, 2021. doi: 10.20535/kpissn.2021.4.251684.

way we can find an estimate for possible loss and construct appropriate risk management algorithms.

Sometimes mathematical models available are too complicated for practical use, and the necessity emerges to create simpler formal descriptions of risk, adequacy of which can be substantially different from the ideal ones. Generally speaking any model represents somewhat simplified representation of the world that can lead to incomplete description of situations with uncertainties, inadequacy and wrong forecasts as well as to incorrect decisions based upon the forecasts. That is why one of the most urgent problems that can be often met in the risk management processes is development of appropriate models adequate enough for practical use. The models should be acceptable for practitioners and, when necessary, supplied with extra structural nonlinear elements; they should have a possibility for correcting prior expert estimates of parameters and initial conditions, restrictions, experimental data, and with clearly defined possible practical applications.

A substantial role regarding timely and high quality problem solving to perform the risk modelling and management plays systemic approach. It means that the approach supposes taking into consideration current market factors (including random ones); possibly revealing new deterministic and stochastic factors of influence (including the hidden ones); estimating the scale and frequency of their influence on the processes of interest; identifying and taking into consideration possible structural, statistical and parametric uncertainties met in the processes of model constructing, estimating of forecasts for relevant processes development and estimating respective risk. In most cases it is necessary to perform correct problem stating and solving optimization problems directed towards risk loss minimization. Another problem related to systemic approach of risk analysis is in providing several sets of statistical quality criteria related to analysis of data quality, adequacy of models, quality of forecasts and decision alternatives generated on the basis of the models and forecasts. These sets of criteria provide the possibility for achieving high quality results of computing at each stage of data processing, model constructing and risk estimation [2], [3]. Usually it is more convenient to perform such analysis in the frames of appropriately designed and implemented specialized decision support system (DSS) [3].

Analysis of financial processes development and their internal and external interaction within the last several decades highlights the needs of special attention, from the point of view of financial risk management, for banking system, investment

and insurance as well as for large and medium level firms. Especially sophisticated and highly dynamic are financial processes in the area of market processes, insurance and respective risk situations. This is explained by the fact that market and insurance areas are directly connected with many other dynamic processes in international banking system, production, tourism, cargo and passenger transportation, and also with natural and industrial catastrophes etc. Just market and actuarial activities are in the focus of solving complex everyday financial problems at all levels of economy and private activities. The activities require correct practical applications of high quality mathematical models, methods and knowledge of data analysis techniques and procedures. The mathematical methods and data models as well as decision support systems based upon them do not replace professionals making final decisions but they provide the possibility for much deeper analysis and understanding of related processes, for improved data and expert estimates processing, for generating possible objective alternatives, and select objectively the best decision for specific application.

A crucial role in analysis of risk and estimating possible loss plays probabilistic approach to data analysis, modelling and forecasting. Say, Bayesian paradigm in the form of systemic Bayesian programming creates appropriate probabilistic-and-statistical instrumentation to fight uncertainties and provide appropriate risk analysis results. Practically all the probabilities we have to cope with in solving most practical problems in every area are conditional. These conditions lead to emerging multiple special cases for analysis of specific problems including risk management. That is why application of the Bayesian modelling approach to risk analysis is very important, useful and appropriate what is supported by many available examples from the past developments.

The study is focused on some aspects of modelling and estimation of financial and some other type of risks, application of related mathematical models to solving practical risk estimation and forecasting problems. The study supposes constructing and implementation of appropriate DSS providing for all necessary computational procedures for reaching high quality results at each step of data analysis, risk modelling and estimation.

Problem statement

The purpose of the study is to: determine basic types of risks in different areas of human activities and consider the possibilities for their mathematical description; determine the possibility of hiring

Bayesian approach to constructing mathematical models of financial as well as some other types of risk, and show examples of constructing Bayesian type models for stochastic financial process.

Operational risks

Definition of many types of risk including the financial one is linked to the probability of events that can be accompanied by some material loss, and the level of the loss. An International Standard Organization gave the following definition to risk: “risk is a combination of probability for some event and its consequences [4]”. The study [5] formulated the following definition: “risk is a set of possible scenarios, s_i , each of which is characterized by the probability p_i , and the consequence c_i ”. This definition is very general, robust and can be hired for solving engineering and financial problems. The most practical problems exhibit multivariate risk, i.e. there exist multiple internal and external risk factors that in combination create general situation leading to emergence of risk.

For example, the actuarial activities are characterized by the set of multiple risks with the most known among them are as follows: individual risks; collective risks for one (short) period of time; collective risks for long periods; high distributed risks of loss; operational risks; non-return risk of credit; bankruptcy risk, and other type of risks [6]–[8]. A substantial loss comes to companies and various enterprises today due to availability of operational risk that exists in any organization and can be considered as the “universal” one. It can be viewed as the risk of direct and indirect loss that comes to being due to inappropriate organization of working activities or inappropriate organization of internal processes in a company, incorrect behaviour of a company staff, and/or incorrect functioning of technical equipment, or due to influence of unfavourable external factors.

The operational risk can also be provoked by the absence of appropriate methods and means for management of this type of risk. The operational risk should be analysed qualitatively and quantitatively as well as any other type of risk that requires collecting and thorough analysing appropriate statistical data and expert estimates. As far as emergence of the risk is influenced by the most different events, the problems of collecting necessary data, model constructing, and estimating the volume of possible loss and its probability require substantial efforts of experts in information technology, mathematical modelling, forecasting and decision support systems.

As a possible source of statistical data could serve insurances policies, that contain information regarding possible insurance risks that come to being due to the events resulting in operational loss. However, this is not the best possibility for collecting necessary information because the policies contain confidential information about clients, and the task of processing policies is complicated, time consuming process that is not distinguished with the information completeness regarding the problem stated. As of today, to get mathematical description for all types of the insurance risks the following methods are widely used: applied statistics, probability theory, fuzzy logic, Bayesian theory of data processing, neural networks etc.

In the process of analysing the situations leading to appearance of financial risk it is important to get objective information regarding current state of insurance company from independent sources. It is necessary to establish correctness of functioning of the risk management department, to study the information accessible for the department and the methods of its processing. One of the most important characteristics of this information is its completeness, i.e. does it contain enough data for constructing the model capable to forecast the volume of possible loss? It may turn out rather often that the information is not complete for discovering all types of possible risks that may emerge in a company. The loss discovered may not reflect all possible types of risks met by the company because not all possible risks resulted in the loss and could be ignored.

In such cases additional analyses should be performed directed towards discovering the following events: (1) determining the moment of time when the financial loss occurred as well as establishing the fact of making the decision resulting in the loss; (2) estimation of the possible income that company could get in the case of avoiding the risky situation; (3) distribution of actual financial loss among the existing several possible risk factors when the number of actual risk factors is greater than one; (4) collection of additional information from the company personnel who is related to the risky situation that took place. The purpose of the additional situational analysis is reconstruction of the sequence of events that actually resulted in the loss, discovering the reasons for happening of the events and establishing the possibilities of their avoiding. It would also be logical to establish the reasons why the events were not avoided. The probability always exists that not all the losses were discovered and were taken into account by the risk management department. That is why the interaction with the company personnel

can reveal additional information regarding other possible loss that was avoided or it remains actual for further studying.

Generally the procedure for identifying the types of risks and their management can be represented in the form of the following cyclical sequence of actions: (1) establishing the possible types of risk for a company; (2) identification, studying, deeper understanding and description of the situations that are favourable regarding development of the risk factors; (3) a thorough analysis of possible risks with establishing appropriate risk measures, loss estimation and forecasting; (4) development and making appropriate control decisions regarding specific risks management; (5) tracking implementation into life the managerial decisions developed, detecting and analysis of the indicators for emergence of possible risks; (6) making detailed reports on the actions directed towards avoiding, ignoring or active control of the situations with risk emergence. Implementation of the cyclical procedure for risk identification and management of respective situations should be based upon classification and continuous analysis of possible risks for a selected company. Very often the term is used like “*enterprise risk*” that includes all possible risks for a specific enterprise. The enterprise risk is divided into *basic risk for business* and *operational risk*; and both types of risk include their specific components.

Today there exist a set of quantitative approaches to risk analysis that can be applied to solving the problem of deeper understanding the essence and estimation the financial risk level. The methods of this class include the following ones: (1) *statistical estimation* – empirical studies; estimation of maximum possible loss; construction of appropriate probability distribution functions; linear and non-linear regression analysis; (2) *frequency analysis of loss* – frequency analysis of the level of loss; theory of extreme values; stochastic differential equations; (3) *statistical Bayesian approach* (Bayesian programming techniques) – dynamic Bayesian models for the systems under study; Bayesian belief networks (BBN), and causal models; process development charts; (4) *artificial intelligence systems* – neural nets, neuro-fuzzy models and decision trees for client and enterprises classification as well as estimation of the possible loss risk; (5) *Monte Carlo based models* and the *mode switching models* – generation of the development scenarios, methods of income analysis; strategic investment analysis; (6) *expert estimation and fuzzy logic* – direct likelihood estimation for the development scenarios; Delphi method; the models for capital and price forming; market risk estimation

in actuarial studies; (7) *practical approaches to risk estimation and management* – stress testing and scenario analysis; industrial and business scenarios; dynamic financial analysis; market beta-comparison of separate companies in the frames of market sectors.

The mathematical models that are hired for the formal description of processes related to risk estimation can be classified as deterministic and stochastic. It is known that such models represent simplified formal pictures of possible consequences for the future uncertain events. The uncertainty is represented in this case with the time of emergence of the events and their consequences, as possible loss that in most cases can be related to the financial one. Though it should be stressed that when mathematical model is hired for a study the forecast estimate can be computed in the form of quite definite value but this “certainty” is based upon the assumptions that are uncertain by their nature. It is useful to remember that the data available practically always contain deterministic and stochastic components, and this fact requires including at least one stochastic variable into deterministic model.

This is related to the probability distributions of random variables used for analysis, selection of the methods for model structure and parameter estimation, determining the type and parameters for stochastic disturbances, selecting the technique for forecast estimation etc. If the assumptions accepted in the process of model constructing correspond to actual behaviour of the process being studied and the future developments of the process correspond to the forecast estimates generated by the model then such estimates could be used for decision making. Probabilistic models are directed to forecasting probabilities of the events (processes) using the historical information regarding former behaviour of the processes. The models are based upon the hypothesis and assumptions that are logically coordinated with the probable development of the events in the future taking into consideration possible uncertainties. It should be noted here that the forecasts computed with the models cannot be completely deterministic (certain) because their complete certainty would actually mean practical uselessness of the model.

Bayesian programming methods

Today there exist a set of Bayesian methods of filtering, modelling, forecasting and decision making known under the general name of Bayesian programming or Bayesian methodology. The methodology provides the possibility for solving the following problems:

– constructing probabilistic and statistical models (model structure and parameters estimation) using statistical (experimental) data and expert estimates;

– computing final results on the basis of the models created according to the specific problem statement: estimates of forecasts; control actions; estimates of variables and parameters using filters; image recognition; making managerial decisions regarding the process and systems under investigation, and many other tasks;

– analysis of quality of the results received at each stage of data processing by making use of appropriate sets of quality criteria.

Some methods related to the Bayesian programming methodology include the techniques in short mentioned below.

1. Recursive Bayesian estimation: filtering, forecasting, and smoothing the data. The basic equation of estimation has the following form:

$$P(S(k) | O(0)...O(k)) = P(O(k) | S(k)) \times \sum_{S(k-1)} \left[P(S(k) | S(k-1)) \times P(S(k-1) | O(0)...O(k-1)) \right],$$

where, $S(0), \dots, S(k)$, is time series of state variables; $O(0), \dots, O(k)$ is time series of observations; $P(S(k) | S(k-1))$ is system or state transition model; $P(O(k) | S(k))$ is conditional model for observations that shows what would be observation at the moment k , if the system is in the state $S(k)$.

This model can be used for computing future distribution of states $P(S(k+l) | O(0)...O(k))$ at the moment $k+l$ on the basis of the available observations $O(0), \dots, O(k)$.

When $l=0$, the filtering procedure is implemented; if $l>0$, then forecast estimate is generated, and when $l<0$, the data is being smoothed. Smoothing means restoring of former state on the basis of the observations received before or after the moment of smoothing.

2. Hidden Markov models (HMM) – is modification of Bayesian filter which supposes that data is discrete; the state transition models and observations are defined by the probability matrices or conditional probability tables. If the variables observed are continuous such models are called semi-continuous HMM.

3. Optimal recursive Kalman filters (KF) can process continuous or discrete data. The state transition and observation models are constructed in this case with the use of Gaussian processes for describing external random disturbances and measurement noise. If nonlinear models are hired then Taylor expansion is used what results in application of local linear models. For simultaneous estimation of

parameters and states an Extended Kalman Filter (EKF) can be applied.

4. Particle (granular) filters (PF) are used for data processing using the following distribution model: $P(S(k-1) | O(0), \dots, O(k-1))$, where $S(\cdot)$ is matrix of states, and $O(\cdot)$ is observation matrix. The observations are approximated by the set of particles with weighting coefficients proportional to the probabilities of their occurring. The state probabilities are renewed by recursive procedure.

5. Static Bayesian networks (BN) are probabilistic and statistical models used to describe formally available data and expert estimates in conditions of uncertainty. To the net variables are imposed practically no restrictions, and no special semantics is used for their description. Thus, definite freedom exists for selecting the variables and constructing the network model.

Graphically BN is represented by directed acyclic graph the vertices of which are variables of the network, and the arcs indicate existing conditional relations between the variables. Formally the model is described by the triple: $BN = \{V, G, T\}$, where V represents the set of variables for constructing the net (data base); G is directed acyclic graph constructed with application of optimization procedures; T is conditional probability tables associated with the graph vertices (for parent vertices these probabilities are unconditional).

The model parameters are created by conditional probabilities in the conditional probability tables. Continuous variables are represented by the specific distributions. If BN contains both discrete and continuous variables it is called a hybrid one. Continuous variables in such cases are transformed into discrete ones using known procedures.

The result of constructing and application of BN is provided in the form probabilistic inference like following: $P(X^i | \text{Known})$, where *Known* means the set of other network variables the probabilities of states for which are known. Generally the probabilistic inference generated by BN is in propagation of probabilities and parameters of Gaussian distribution laws over the whole network and depends on availability of evidences – additional information about the network states. The process of forming (computing) the inference is based upon rather complicated mathematical algorithms. One of the simplest approaches is based upon Bayes rule.

6. Dynamic Bayesian networks (DBN) are models of the same type as static BN but the difference is that they take into account dynamics of the processes being modelled (i.e. their evolution in time) as well as possible stochastic influences to the

basic variables. Actually, DBN represents further evolution of static networks directed towards taking into consideration evolution of the processes considered in time. First static network is constructed for the variables selected and data available, that is considered as time invariant one. Then this model structure is repeated for each next moment of time with coming new observations (measurements).

7. Part of the model graph that corresponds to the specific moment of time t_k , or simply k , is called time cross-section. If hypothesis is accepted that the current system state (being modelled) depends on the previous state only then this assumption is called *first order Markov assumption*. And if the model structure remains unchanged for all the time cross-sections then such DBN is called stationary one. In such cases model that corresponds to one moment of time is defined as local and time-invariant or homogeneous.

8. Markov localization (ML) models: these are models of the Bayesian filter type that additionally include control variables, $\mathbf{u}(0), \dots, \mathbf{u}(k-1)$. They also have another name of *hidden Markov input-output models*. Such models provide a possibility for improving state estimates using control variables as follows: $P(S(k) | \mathbf{u}(k-1), S(k-1))$. Sometimes the model is called action model. Generally these models are represented in matrix form similar to those used in particle filters under the name of Monte Carlo with Markov localization (MCML). The model constructed should give an answer to the question: what is the probability of current state of a system under study:

$$P(S(k) | \mathbf{u}(0), \dots, \mathbf{u}(k-1), O(0), \dots, O(k)),$$

using the information regarding former control actions and observations of previous states. The term "localization" is linked to robotic system applications, i.e. localization of a robot in space. The basic state equation is similar to the basic filtering equation given above.

9. *Bayesian maps*: the name that means generalization of the Markov localization models that also were originated in the sphere of robotic control systems.

10. Bayesian approach to data processing and decision making using hierarchical models.

11. Bayesian regression in the form of generalized linear models (GLM).

Bayesian approach to modelling and decision making

The Bayesian approach to modelling and computing probabilistic inference supposes fulfilling the following stages given in Fig. 1. According to the approach the initial information is coming from two sources: prior information from researcher and statistical data received as a result of performing appropriate experiments.

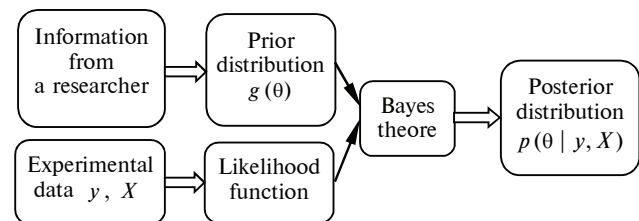


Fig. 1 Information flows in the data processing system based upon Bayes theorem

The prior information from a researcher is based upon the results of his former studies; theoretical information regarding the processes under study; informal data received from various sources. Generally this is extra information that is added to the experimental one. The Bayes rule (BR) is in the center of the methodology, and it can be written as follows:

$$p(\theta | y, \mathbf{X}) = \frac{g(\theta)p(y | \theta, \mathbf{X})}{p(y)} \mu g(\theta)p(y | \theta, \mathbf{X}), \quad (1)$$

where y is observation vector for dependent variable for the system under study; \mathbf{X} is observation matrix for independent variables (regressors) that influence behaviour of the dependent one; θ is vector of random parameters that together with matrix \mathbf{X} define probability density function (PDF) for y (if we refer to classic regression, θ is vector of the regression); $p(\cdot)$ is PDF; $p(\theta)$ is prior density for random parameter θ , that is based upon prior knowledge of a researcher (y, \mathbf{X}) before applying experimental results; $p(\theta | y, \mathbf{X})$ is conditional PDF for the data y given θ and \mathbf{X} , in other words, this is likelihood function for data y ; $p(y)$ is marginal probability for y after excluding influence of independent variables and parameters. Here $p(\theta | y, \mathbf{X})$ is posterior PDF for parameter, θ , based upon renovation of the prior distribution by the data (y, \mathbf{X}).

Thus BR shows the way how to combine prior distribution with experimental data to correct the prior PDF. Actually the BR provides the possibility

not only for combining information coming from two sources but any other way of combining will violate logical (and mathematical) essence of the rules for operating with PDF.

The expression (1) contains two variants of representing the Bayes rule. The first one is given in the form of equality due to availability of denominator $p(y)$ (unconditional density for, y) that plays the role of normalizing constant providing condition that posterior conditional density for θ is proper and integrates to 1 over the parameter definition range. The second variant of the rule representation is given without normalizing constant what simplifies the expression. In practice, first the nominator is computed, and then the normalizing constant if necessary. In some cases only ratio of results is of interest and the normalizing constant is cancelled.

Application of Bayesian approach to risk modelling and estimation has the following advantages:

- this approach is convenient for taking into consideration various uncertainties that are related to statistical data like imputation of missing observations, noise and short samples processing; appropriate state and parameter estimation procedures; application of different data formats; hiring of expert estimates in modelling procedures; possibilities for application of high dimension of data and models; the use of static and dynamic approaches to modelling various systems; availability of alternative procedures for model structure and parameter estimation; possibilities for model adaptation and constructing simulation procedures etc.;

- possibilities for constructing complex hierarchical models;

- possibilities for constructing combined models that include linear and nonlinear regression equations, Bayesian and neural networks, fuzzy logic etc.;

- taking into consideration probabilistic uncertainties of the type: “will some event happen or not”, and “under what conditions it will happen”;

- elimination of amplitude-like uncertainties by replacing missing measurements with estimated or generated values;

- availability of linear and nonlinear filtering procedures;

- possibilities for application of forecasting distributions;

- availability of possibilities for application of Bayesian optimization.

Bayesian approach to the formal description of risk

Consider possible approach to formal description of a risk using Bayesian approach. Let $\mathbf{M} = \{M_1, M_2, \dots, M_n\}$ is a set of risk describing models; $p(x | M_i, \theta_i)$ is a likelihood function for the model $M_i, i=1, \dots, N$ with parameters θ_i and available data, x ; and $p(\theta_i | M_i)$ is prior distribution for parameter vector θ_i related to selected model M_i . According to Bayes rule posterior distribution of vector θ_i given known model structure M_i and data x can be written in the form:

$$p(\theta_i | M_i, x) = c_i p(x | M_i, \theta_i) p(\theta_i | M_i),$$

where c_i is normalizing constant. Various model structures may differ with the number of parameters and better models are considered those that have lower number of parameters, according to the economy principle, but provide the possibility for acceptable quality (adequacy) of formal description for the process under study. To select the best model it is possible to use appropriately modified Akaike information criterion (AIC) [9]:

$$\log \tilde{p}(x | M_i, \theta_i) = \log(x | M_i, \theta_i) - A(k_i),$$

where k_i is dimension of vector θ_i ; $A(k_i) = k_i$ is increasing function of k_i . This criterion can be represented in the form that is minimized for the best model, i.e. $-2 \log p(x | M_i, \theta_i) + 2k_i$. An alternative for the AIC is Bayes-Schwartz criterion, where $A(k_i) = 0.5k_i \log N$, and where N is sample power for data $\{x\}$. Some other modifications of the criteria are hired for analysis of regression models of various structures.

After selection of the model structure, $M^* = M_{i^*}$, it is necessary maximize posterior density $p(\theta_{i^*} | M_{i^*}, x)$ for θ_{i^*} to find better estimate of the parameter vector $\hat{\theta}_{i^*}$. If the power of sample $\{x\}$ is large enough, and the prior distribution $p(\theta_{i^*} | M_{i^*})$ is of diffusion type then the posterior maximum can be replaced by the maximum likelihood estimate. The model structure, M_{i^*} , and parameters, θ_{i^*} , found this way can be considered as acceptable ones. Now consider the possibility of introducing uncertainties into the model and parameters.

Let $p_r(M)$ is prior probability for estimating model structure M , and $p(\theta | M_i)$ be conditional prior distribution for parameters θ with known structure M . According to the Bayes rule it can be written:

$$p(\theta, M | x) = p(\theta | m, x)p_r(M | x) = p(\theta | m, x) \times c \times p_r(M)p(x | M),$$

where $p(x | M) = \int p(x | \theta, M)p(\theta | M)d\theta$; and c is normalizing constant. Thus, joint posterior density for (θ, M) is determined by the product of posterior density for θ , with condition that model M is adequate, and posterior probability of determining adequate structure M for the data $\{x\}$:

$$p(\theta, M | x) = c(c_i p(x | \theta, M)p(\theta | M)) \times p(x | M)p_r(M),$$

where c_i is normalizing constant for $p(\theta | M, x)$, and c is normalizing constant for the whole posterior distribution.

Suppose it is necessary to determine the probability of bankruptcy, y , that is determined via parameters of model M . Conditional probability for y can be written as follows:

$$p(y | x) = \sum_i p(y | x, M_i)p_r(M_i | x),$$

where $p(y | x, M_i) = \int p(y | x, M_i, \theta_i)p(\theta_i | M_i, x)d\theta_i$;

$$p(\theta_i | M_i, x) = c_i p(x | M_i, \theta_i) p(\theta_i | M_i).$$

An asymptotic analysis of the conditional distribution $p(\theta_i | M_i, x)$ says [10], [11], that

$$\log p(\theta_i | M_i, x) = c - \frac{1}{2}(\theta_i - \hat{\theta}_i)^T H_i(\theta_i - \hat{\theta}_i) + o(|\theta_i - \hat{\theta}_i|^2),$$

where $\hat{\theta}_i$ is weighted mean of maximum likelihood estimate for θ_i , and of the prior distribution mode for θ_i ; H_i is a sum of Hessians in corresponding maxima of likelihood function and prior density.

If the prior distribution is of diffusion type and power of sample N is large enough then parameter estimate computing can be performed using appropriate approximation. For example, matrix H_i can be approximated by the product, $N\hat{B}_i$, where $\hat{B}_i = B_i(\hat{\theta}_i)$ is information matrix for one observation with condition that $\hat{\theta}_i$ is actual value for θ_i . The estimated $\hat{\theta}_i$ can also be computed with maximum likelihood technique.

Now consider posterior density for the model M_i :

$$p_r(M_i | x) = cp_r(M_i)p(x | M_i),$$

where $p(x | M_1), p(x | M_2), \dots$ are Bayes factors approximated by the scaling constant. Somewhat sim-

plified computing of the factors can be performed using the expression [12], [17]:

$$\log p(x | M_i) = c + \frac{1}{2}k_i \log 2\pi - \frac{1}{2} \log |\hat{I}_i| + \log p(x | \hat{\theta}_i, M_i) + \log p(\hat{\theta}_i | M_i) + O(N^{-1}), \quad (1)$$

where \hat{I}_i is information matrix for the data $\{x\}$, that is described by the model having parameters $\hat{\theta}_i$; k_i is dimension of the parameter vector $\hat{\theta}_i$; $p(\theta_i | M_i)$ is a prior distribution for $\hat{\theta}_i$. As far as the observations are independent as assumed, then, $\hat{I}_i = N\hat{B}_i$, where \hat{B}_i is information matrix that corresponds to a single observation when the model M_i with parameters $\hat{\theta}_i$ is used. Thus, it can be written:

$$\log |\hat{I}_i| = k_i \log N + \log |\hat{B}_i|.$$

It should be noted here that with growing N the second member in the RHS will be approximately constant for each model. In a case when k_i accepts the same value for each model the only $\log |\hat{B}_i|$ will be varying. An influence of the member $\log p(\hat{\theta}_i | M_i)$ that is linked to prior distribution is negligible, especially when the diffusion distribution is used. After introducing the notation of $\hat{l}_i = \log p(x | \hat{\theta}_i, M_i)$, and eliminating the member $\log p(x | \hat{\theta}_i, M_i)$, the criterion (1) is transformed into the following simplified form:

$$\log p(x | M_i) \approx c + \frac{1}{2}k_i \log 2\pi + \hat{l}_i - \frac{1}{2}k_i \log N - \frac{1}{2} \log |\hat{B}_i|,$$

or

$$\log p_r(M_i | x) \approx \log p_r(M_i) + \frac{1}{2}k_i \log 2\pi + \hat{l}_i - \frac{1}{2}k_i \log N - \frac{1}{2} \log |\hat{B}_i| + c,$$

Where c is a normalizing constant that provides for holding the equality: $\sum_i p_r(M_i | x) = 1$. Now consider examples of application of the Bayesian approach to financial risk estimation using appropriate statistical data.

Example 1. The model of randomly incoming payments. Let $\{x(k)\}$ is stochastic process of incoming payments, where k denotes discrete moments of time. Thus the total sum of payments for the first moment of time can be represented as $\exp(x(1))$,

and for arbitrary moment k the stored capital will amount to $\exp(x(1) + x(2) + \dots + x(k))$. For the sake of convenience denote the sum under exponent as follows: $y(k) = \sum_{i=1}^k x(i)$. It is necessary to determine the types of distributions for $y(k)$, and $F(k) = \exp(y(k))$. One of the simplest models that could be hired to formally describe the financial process is low order autoregression, say AR(1):

$$x(k) = a_0 + a_1 x(k-1) + \varepsilon(k), \quad (2)$$

where $\varepsilon(k)$ is stochastic process including random disturbances, measurement errors as well as model structure inadequacies. Usually the process is considered as a normal one due to growing number of measurements and multiple ingredients of the noise. Rewrite equation (2) in a more convenient for further analysis form:

$$x(k) - \mu = a(x(k-1) - \mu) + \sigma z(k), \quad (3)$$

where μ is sample mean for the data $\{x(k)\}$; a, σ are model parameters; $\{z(k)\} \sim \text{Norm}(0,1)$ is a standard normal distribution sequence. Now find expressions for the model parameter estimates using maximum likelihood technique.

The conditional likelihood function for the sequence of payments $\{x(k)\}$ is as follows:

$$f(x | \mu, \sigma^2, a) = \prod_{i=-N+2}^0 \left\{ (2\pi\sigma^2)^{-\frac{1}{2}} \times \exp \left[-\frac{(x(i) - \mu - a(x(i-1) - \mu))^2}{2\sigma^2} \right] \right\}. \quad (4)$$

Using the function the following expressions for model (3) parameter estimates can be written as follows:

$$\hat{a} = \frac{\sum_{i=-N+2}^0 (x(i) - \hat{\mu})(x(i-1) - \hat{\mu})}{\sum_{i=-N+2}^0 (x(i-1) - \hat{\mu})^2};$$

$$\hat{\mu} = \frac{1}{N-1} \left(\sum_{i=-N+1}^{-1} x(i) + \frac{x(0) - x(-N+1)}{1 - \hat{a}} \right);$$

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{i=-N+2}^0 [x(i) - \hat{\mu} - \hat{a}(x(i-1) - \hat{\mu})]^2.$$

Now write approximate expression for the conditional likelihood function (4) using the Taylor expansion:

$$f(x | \mu, \sigma^2, a) \propto (\sigma^2)^{-\frac{N-1}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \left[\begin{aligned} &\varphi_1 + \varphi_2(\mu - \hat{\mu})^2 + \\ &+ \varphi_3(a - \hat{a})^2 + \\ &+ \varphi_4(\mu - \hat{\mu})(a - \hat{a}) + \\ &+ \varphi_5(\mu - \hat{\mu})(a - \hat{a})^2 + \\ &+ \varphi_6(\mu - \hat{\mu})^2(a - \hat{a}) + \\ &+ \varphi_7(\mu - \hat{\mu})^2(a - \hat{a})^2 \end{aligned} \right] \right\},$$

where " \propto " stands for proportionality; $\varphi_1 = (N-1)\hat{\sigma}^2$; $\varphi_2 = (N-1)(1-\hat{a})^2$; $\varphi_3 \approx (N-1)\hat{\sigma}^2 / (1-\hat{a}^2)$; $\varphi_4 = 2(x(-N+1) - x(0)) \approx 0$; $\varphi_5 = \varphi_4 / (1-\hat{a}) \approx 0$; $\varphi_6 = -2(N-1)(1-\hat{a})$; $\varphi_7 = (N-1)$.

To solve the problem of modelling risk in the problem statement given above the following prior parameter distribution is hired [13], [14]:

$$f(\mu, \sigma^2, a) = \sigma^{-3}(1-a)^{1/2}(1+a)^{-1/2},$$

or $f(\mu, \sigma^2, a) = \sigma^{-3}(1-a)^{3/2}(1+a)^{1/2},$

that can be used when $a \neq \pm 1$. Now the conditional posterior distribution for the process parameters can be written as follows:

$$(\mu | x, \sigma^2, a) \sim \text{Norm} \left(\hat{\mu}, \frac{\sigma^2 / (N-1)}{(1-a)^2} \right);$$

$$f(a | x) = d(\hat{a}, N) [1 - \hat{a}^2 + (a - \hat{a})^2]^{-\frac{N-1}{2}} (1-a^2)^{\frac{1}{2}},$$

$$-1 < a < 1.$$

Now, taking into consideration the expressions for parameter μ, σ^2, a distributions given data, $\{x(k)\}$, return to equation (3):

$$x(i) - \mu = a(x(i-1) - \mu) + \sigma z(i) = \sigma(z(i) + az(i-1) + \dots + a^{s-1}z(1)) + a^s(x(0) - \mu).$$

The results achieved provide the possibility for writing expression for the process $y(k)$ in the form:

$$y(k) | \mu, \sigma^2, a, x = \mu k + (x(0) - \mu)M(a, k) + \sigma(V(a, k))^{\frac{1}{2}}Z,$$

where, $M(a, k) = a(1 - a^k) / (1 - a)$;

$$V(a, k) = \frac{1}{(1-a)^2} \left(k - \frac{2a(1-a^k)}{1-a} + \frac{a^2(1-a^{2k})}{1-a^2} \right).$$

To perform computational experiment the data regarding incoming payments for a selected insur-

ance company was gathered within five year period. As a result the following parameter estimates were found:

$$\hat{a} = -0.2587; \hat{\mu} = 0.2165; \hat{\sigma}^2 = 0.0873.$$

The prior probability for the model was selected equal to 0.5, what is quite logical in conditions of absence of extra information regarding the estimate. The posterior probability for the model let to the value of, $p_r(M) = 0.59$. Thus the probabilistic approach demonstrated its advantage: the parameters of forecasting posterior distribution in conditions of parametric and statistical uncertainties correspond to actual ones. This way we demonstrated how parametric uncertainty could be taken into account.

It is clear that achieved posterior probability for the model generally is not high enough what could be explained by the use of simplified initial model AR(1) and approximate computing of the conditional likelihood function. Also, it is necessary to perform additional study of influence of prior distribution on the final result. The additional deeper study of the possibility for application of probabilistic models of the type considered requires application of several alternative techniques to model building with subsequent comparison of the results achieved for each of them.

Example 2. Quality analysis of automatized client service (operational risk analysis): the case of discrete data and discrete parameters. Insurance company introduced automatized service for clients that provides for automatic registration of clients insuring their cars. The number of clients, users of the service, reaches several thousand per month. Consider the problem of estimating operational error θ after servicing n clients.

To simplify the problem statement suppose that θ can accept the following three values: 0.25 is good result; 0.50 is acceptable result; 0.75 is bad result that cannot be accepted. Available statistic shows that within former two years the company provided the automatized client service with the following quality:

- within 60% of time the probability of service error was at the level of $\theta = 0.25$ (good result);
- within 30% of time the probability of service error was at the level of $\theta = 0.50$ (acceptable result);
- within 10% of time the probability of service error was at the level of $\theta = 0.75$ (result that cannot be accepted).

These results were used as prior probabilities so that to forecast the level of service in the future. This distribution is given in the table 1.

Table 1. Prior probabilities for parameter θ

	Service quality		
	good	acceptable	unacceptable
Probability of service errors (θ)	0.25	0.50	0.75
Density for the errors $\theta(p(\theta))$	0.60	0.30	0.10

After 10000 cases of client services the company decided to perform the service quality control. The control showed that out of ten randomly selected service cases two of them contained errors. What conclusion regarding the service quality should be made in this case? I.e., in other words, what is actual posterior distribution for the parameter, θ ?

In this case the data has discrete form and first it is necessary to determine (say, on the bases of previous experience) type of distribution for data. On the basis of former experience we can suppose that it has binomial distribution with parameter, θ :

$$f(\theta, n, r) = \binom{n}{r} \theta^r (1 - \theta)^{n-r} = C_n^r \theta^r (1 - \theta)^{n-r},$$

where, $\binom{n}{r} = \frac{n!}{r!(n-r)!}$. The number of successful events is equal in this case to $r = 2$; successful are events linked to emergence of errors in the client service out of 10 possible, i.e. $n = 10$. Thus, likelihood function for data in this case has the following form:

$$L(\theta) = \binom{n}{r} \theta^r (1 - \theta)^{n-r} = C_{10}^2 \theta^2 (1 - \theta)^{10-2},$$

where $\theta = [0.25; 0.5; 0.75]$ is distribution of possible events linked to the number of service errors. The nominator of the Bayes rule in this case is as follows:

$$h(\theta | r, n) \propto L(r | \theta, n)g(\theta).$$

Now compute necessary likelihood values, $L(\theta)$:

- if $\theta = 0.25$, then

$$L(\theta) = \frac{n!}{r!(n-r)!} (0.25)^2 (1 - 0.25)^8 = 0.30300;$$

– if $\theta = 0.50$, then

$$L(\theta) = \frac{n!}{r!(n-r)!} (0.50)^2 (1-0.50)^8 = 0.04405;$$

– if $\theta = 0.75$, then

$$L(\theta) = \frac{n!}{r!(n-r)!} (0.75)^2 (1-0.75)^8 = 0.00039.$$

The prior and posterior densities for h are given in table 2.

Table 2. Posterior probabilities

θ	Prior probabilities for θ	Likelihood $L(\theta)$	$h =$ posterior density for θ
0.25	0.60	0.30300	0.955
0.50	0.30	0.04405	0.034
0.75	0.10	0.00039	0.010
Total sum:			0.999

Table 2 shows that the most probable value of θ , estimated on the basis of analysis of service quality in the case of automated servicing is the value: $\theta = 0.25$. It means that quality of service is on the “good” level (posterior probability for θ is 0.955). Before making final conclusion regarding quality of client service using automatized servicing system it is necessary to analyse larger number of cases than, 10; say 500 or 1000.

Conclusions

Thus there exists a wide sector of financial and other type of risks that require analytical studies with application of mathematical, statistical and probabilistic models of various structures and complexity. Some methods of quantitative analysis were determined that can be used for solving the problem of thorough understanding of the origin and estimating the level of possible financial loss. For estimating the level and probability of loss the probabilistic type of models can be successfully hired thanks to the fact they provide the possibility for taking into consideration parametric and statistical uncertainties of the processes under study. Here we considered a spe-

cial case of Bayesian approach to formal description of financial risks and the procedure was proposed for probabilistic model constructing directed towards forecasting the incoming payments. It should be noted that Bayesian approach to risk analysis is a natural way for their formalization and receiving highly useful results of estimation and forecasting possible loss as well as its probability.

The computational experiments were performed aiming at parameter estimation of forecasting distribution. The results achieved are very close to the ones computed by the method of moments applied to actual statistical data. It was estimated that posterior probability for the model created achieved the following value: $p_r(M) = 0.59$. The achieved posterior probability for the model generally is not high enough what could be explained by the usage of simplified initial model AR(1) and approximate computing of the conditional likelihood function.

Another example of the Bayesian approach application showed that Bayes rule can be applied directly to solving the problem of analysing risky situations leading to emergence of operational risk. In this case we estimated the state of automatized servicing system and the risk of erroneous service that may result in operational loss. Thus, general conclusion regarding possible application of the Bayesian approach to risk estimation is positive what coincides with numerous other available examples of application the theory.

In the future studies it will be necessary to apply a set of alternative probabilistic candidate models aiming selection of the best one for high quality forecasting of possible loss. To estimate effective forecasting models it will be perspective to apply Monte Carlo procedures for Markov chains. In this case Markov chains will represent unknown model parameter estimates. Such numerical approach to model constructing provides a possibility for substantial increase of a number of parameters and to develop more complex model structures. Application of the Monte Carlo method will also provide the possibility for constructing simulation models that will explain away many hidden interactions of the risk factors involved.

References

- [1] P.L. Bernstein, *Against the Gods: The Remarkable Story of Risk*, NY, USA, Wiley, 1996.
- [2] A.J. McNeil, R. Frey and P. Embrechts, *Quantitative Risk Management*. Princeton, NJ, USA, Princeton Univ. Press, 2005.
- [3] C.W. Holsapple and A.B. Winston, *Decision Support Systems*. Eagan, MN, USA: West Publishing Company, 1996.
- [4] T. Aven and O. Renn, “On risk defined as an event where the outcome is uncertain”, *Journal of Risk Research*, no. 12, pp. 1–11, 2009.

- [5] S. Kaplan, "The words of risk analysis", *Risk Analysis*, vol. 17, pp. 407–417, 1997.
- [6] N.L. Bowers, H.U. Gerber, J.C. Hickman, D.A. Jones and C.J. Nesbitt, *Actuarial Mathematics*. Itaxa, IL, USA, The Society of Actuaries, 1986.
- [7] P. De Jong and G.Z. Heller, *Generalized Linear Models for Insurance Data*. Cambridge, MA, USA: Cambridge Univ. Press, 2008.
- [8] Ph. Jorion, *Financial risk Manager Handbook*. NY, USA, Wiley, 2003.
- [9] Sik-Yum Lee, *Structural Equation Modeling*. NY, USA, Wiley, 2007.
- [10] J.M. Bernardo and A.F.M. Smith, *Bayesian Theory*. NY, USA, Wiley, 2001.
- [11] A. Gelman, J.B. Carlin, H.S. Stern and D.B. Rubin, *Bayesian Data Analysis*. Boca Raton, FL, USA, Chapman and Hall/CRC, 2004.
- [12] D. Draper, "Assessment and propagation of model uncertainty", *Journal of the Royal Statistical Society, Ser. B*, vol. 57, pp. 45–97, 1995.
- [13] S. Klugman, *Bayesian Statistics in Actuarial Science*. Amsterdam, NL: Kluwer, 1992.
- [14] P.E. Rossi, G.M. Allenby and R. McCulloch, *Bayesian Statistics and Marketing*. NY, USA: Wiley, 2005.
- [15] C.P. Robert, *The Bayesian Choice*, NY, USA: Springer, 2007.
- [16] N.V. Kuznietsova and P.I. Bidyuk, *Theory and Practice of Analysis Financial Risks*. Kyiv, UA: Lyra-K, 2020.
- [17] M.Z. Zgurovsky, P.I. Bidyuk, O.M. Terentiev, and Prosyankina-Zharova, *Bayesian Networks in Decision Support Systems*. Charkov, UA: Edelweis, 2015.

Н.В. Кузнєцова, О.М. Трофимчук, П.І. Бідюк, О.М. Терентьєв, Л.Б. Левенчук

ЙМОВІРНІСНЕ МОДЕЛЮВАННЯ РИЗИКІВ РІЗНОЇ ПРИРОДИ

Проблематика. Усім видам людської діяльності притаманні певні ризики, зокрема фінансові. Відповідно, існує проблема оцінювання і прогнозування можливих втрат, для розв'язання якої необхідно створити адекватний математичний опис для формального представлення обраних ризиків. Оцінювання можливих втрат може ґрунтуватись на обробленні наявних даних та експертних оцінок, що характеризують історію та поточний стан процесів, які аналізуються. Належний інструментарій для моделювання та оцінювання ризиків можливих втрат забезпечується використанням ймовірнісного підходу, який включає баєсові методи, відомі на сьогодні як методологія баєсового програмування.

Мета дослідження. Зробити огляд деяких методів баєсового аналізу даних, які забезпечують можливість побудови моделей вибраних ризиків. Зокрема, використати статистичні дані для формального опису операційного ризику, що може з'явитись у процедурах обробки інформації.

Методика реалізації. Для обробки даних і побудови моделей використовується методологія баєсового програмування. Для оцінювання операційного ризику також застосовано теорему Баєса у формулюванні для дискретних подій та дискретних параметрів.

Результати дослідження. На основі запропонованого підходу побудовано модель операційного ризику, пов'язаного з некоректною обробкою інформації. Для того, щоб побудувати і застосувати модель для оцінювання ризику, проаналізовано задачу оцінювання ризику, вибрано змінні та оцінено умовні апіорні ймовірності. Функціонування побудованих моделей продемонстровано на ілюстративних прикладах.

Висновки. Практично важлива задача моделювання та оцінювання ризиків різних типів, зокрема фінансових, може бути вирішена методами баєсового програмування, які дають можливість ідентифікувати і врахувати невизначеності даних та експертних оцінок. Модель ризику, побудована за допомогою запропонованого методу, ілюструє можливості застосування баєсових методів для розв'язання задачі оцінювання ризиків.

Ключові слова: фінансові процеси; фінансові ризики; методологія баєсового програмування; оцінювання ризику.

Рекомендована Радою
Навчально-наукового інституту прикладного системного аналізу
КПІ ім. Ігоря Сікорського

Надійшла до редакції
18 листопада 2020

Прийнята до публікації
14 лютого 2022 року

DOI: 10.20535/kpissn.2021.4.242769

UDC 519.832

V.V. Romanuke

Polish Naval Academy, Gdynia, Poland

romanukevadimv@gmail.com

FINITE APPROXIMATION OF ZERO-SUM GAMES PLAYED IN STAIRCASE-FUNCTION CONTINUOUS SPACES

Background. There is a known method of approximating continuous zero-sum games, wherein an approximate solution is considered acceptable if it changes minimally by changing the sampling step minimally. However, the method cannot be applied straightforwardly to a zero-sum game played with staircase-function strategies. Besides, the independence of the player's sampling step selection should be taken into account.

Objective. The objective is to develop a method of finite approximation of zero-sum games played in staircase-function continuous spaces by taking into account that the players are likely to independently sample their pure strategy sets.

Methods. To achieve the said objective, a zero-sum game, in which the players' strategies are staircase functions of time, is formalized. In such a game, the set of the player's pure strategies is a continuum of staircase functions of time, and the time is thought of as it is discrete. The conditions of sampling the set of possible values of the player's pure strategy are stated so that the game becomes defined on a product of staircase-function finite spaces. In general, the sampling step is different at each player and the distribution of the sampled points (function-strategy values) is non-uniform.

Results. A method of finite approximation of zero-sum games played in staircase-function continuous spaces is presented. The method consists in irregularly sampling the player's pure strategy value set, solving smaller-sized matrix games, each defined on a subinterval where the pure strategy value is constant, and stacking their solutions if they are consistent. The stack of the smaller-sized matrix game solutions is an approximate solution to the initial staircase game. The (weak) consistency of the approximate solution is studied by how much the payoff and optimal situation change as the sampling density minimally increases by the three ways of the sampling increment: only the first player's increment, only the second player's increment, both the players' increment. The consistency is decomposed into the payoff, optimal strategy support cardinality, optimal strategy sampling density, and support probability consistency. It is practically reasonable to consider a relaxed payoff consistency.

Conclusions. The suggested method of finite approximation of staircase zero-sum games consists in the independent samplings, solving smaller-sized matrix games in a reasonable time span, and stacking their solutions if they are consistent. The finite approximation is regarded appropriate if at least the respective approximate (stacked) solution is ε -payoff consistent.

Keywords: game theory; payoff functional; staircase-function strategy; matrix game; irregular sampling; approximate solution consistency.

Introduction

Zero-sum (also known as antagonistic) games are usually used to model processes where two sides referred to as persons or players interact in struggling for optimizing the to-be-paid-or-pay events [1], [2]. A possible action of the player called its (pure) strategy can be intended to bring a particular contribution into the interaction process in order to receive the best payoff for this player [3, 4]. The strategy can be as a simple (point) action whose duration is usually short, as well as a process consisting of an order of simple actions [3, 5].

The simplest zero-sum game is a matrix game, whichever the pure strategy complexity is. Any matrix game has optimal solutions – one, a finite number, or continuum, either in pure or mixed strategies [1, 6, 7]. A more complicated zero-sum game is the antagonistic game, in which the game payoff function (kernel) is a surface defined on a finite-dimensional compact Euclidean subspace. A simple example of the subspace is a unit square [1, 6, 7]. In such cases, opposed to matrix games, the optimal solution is not always determinable. Additional complications may arise when the surface has discontinuities [8, 9]. Moreover, zero-sum games

Рекомендуємо цитувати цю статтю так: V.V. Romanuke, "Finite approximation of zero-sum games played in staircase-function continuous spaces", *Наукові вісники КНІІ*, № 4, с. 19–38, 2021. doi: 10.20535/kpissn.2021.4.242769.

Please cite this article as: V.V. Romanuke, "Finite approximation of zero-sum games played in staircase-function continuous spaces", *KPI Science News*, no. 4, pp. 19–38, 2021. doi: 10.20535/kpissn.2021.4.242769.

defined on open or half-open subspaces (e. g., open square like in the examples from [6, 7]) may not have an optimal solution at all [1, 5, 6]. Therefore, aiming at assuredly obtaining an optimal solution, rendering a zero-sum game to a matrix one becomes a crucial task in zero-sum-game modeling. Without rendering, a zero-sum game may have an intractable optimal solution (if any), when the optimal strategy support is infinite or continuous (e. g., see the examples in [6, 7, 10, 11]). A zero-sum game, in which the player's strategy is a function (e. g., of time), is a far more complicated case. In such games, the payoff kernel must be a functional mapping every pair of functions (pure strategies of the players) into a real value [7, 8, 12, 13]. A game played with such function-strategies is rendered down to a matrix game only when each of the players possesses a finite set of one's function-strategies. Obviously, the rendering is theoretically impossible if the set of the player's strategies is infinite.

The question of rendering an infinite game to a finite one was studied in [14, 15]. The core consists in approximating the infinite game so that the approximated game would not lose the properties of the initial game. There are two fundamental conditions in the game approximation core that allow rendering a zero-sum game with strategies as functions down to a matrix game. First, a time interval, on which the pure strategy is defined, should be broken into a set of subintervals, on which the strategy could be (maybe, approximately) considered constant. The second condition requires that the set of possible values of the player's function-strategy be finite.

The first fundamental condition is the time sampling condition. It can be done according to the rules of a system to be game-modelled, where the administrator (supervisor, manager, controller, etc.) does always define (or constrain) the form of the strategies players will use [8, 10, 16, 17]. Moreover, any process is interpreted static on a sufficiently short time span. Henceforward, the time sampling condition is considered fulfilled (automatically, by default). Then the function-strategy becomes staircase. To keep the terminology simple, the respective game can be called staircase.

The second fundamental condition is imposed for the natural reason that the number of factual actions of the players (in any game) is always finite. While the players may use strategies of whichever form they want, the number of their actions has a natural limit (unless the game is everlasting; but the everlasting game is an unreal mathematical object) [6, 7, 10, 17]. Thus, the set of function-strategies

used in a zero-sum game is finite anyway. Therefore, any non-everlasting zero-sum game is played as if it is a matrix game. However, the size of this matrix game depends on how each of the players has decided on discretizing (i. e., finitely approximating) one's set of function-strategy values. It does not seem that a player is likely to independently discretize the set identically to the other player's discretization.

Theoretically, the continuous game approximation is based on sampling (discretizing) either the payoff kernel or the sets of players' pure strategies. Basically, this is the same. A method of approximating continuous zero-sum games is known from [14, 15, 18]. An approximate solution is considered acceptable if it changes minimally by changing the sampling step minimally. Obviously, the independence of the player's sampling step selection should be taken into consideration. Moreover, the method cannot be applied straightforwardly to a zero-sum game played with staircase-function strategies. However, a part of the staircase game considered on a time subinterval where the players' strategies are constant can be directly approximated by the method taking into account the independence of the player's sampling step selection.

Problem statement

Issued from the impossibility of solving zero-sum games played in staircase-function continuous spaces, the objective is to develop a method of finite approximation of such games by taking into account that the players are likely to independently sample their pure strategy sets. For achieving the objective, the following six tasks are to be fulfilled:

1. To formalize a zero-sum game, in which the players' strategies are functions of time.

2. To formalize a zero-sum game, in which the players' strategies are staircase functions. In such a game, the set of the player's pure strategies is a continuum of staircase functions of time, and the time is thought of as it is discrete.

3. To state conditions of sampling the set of possible values of the player's pure strategy so that the game be defined on a product of staircase-function finite spaces. By this, the sampling step is to be different at each player. In addition, the distribution of the sampled points (function-strategy values) must not be necessarily uniform.

4. To state conditions of the appropriate finite approximation. This implies also convergence. The independence of the player's sampling step selection is to be discussed also.

5. To discuss applicability and significance of the method for the game theory. In particular, the question of how to reconcile the difference of the players' sampling step selection is to be discussed as well.

6. Make an unbiased conclusion with a clear prompt of how the research might be extended and advanced.

A zero-sum game played with strategies as functions

A zero-sum game, in which the player's pure strategy is a function of time, is defined by the following presumptions. Let each of the players use time-varying strategies defined almost everywhere on interval $[t_1; t_2]$ by $t_2 > t_1$. The first player's strategy is denoted by $x(t)$ and the second player's strategy is denoted by $y(t)$. These functions are bounded, i. e.

$$a_{\min} \leq x(t) \leq a_{\max} \text{ by } a_{\min} < a_{\max} \quad (1)$$

and

$$b_{\min} \leq y(t) \leq b_{\max} \text{ by } b_{\min} < b_{\max}. \quad (2)$$

Besides, the square of the function-strategy is presumed to be Lebesgue-integrable [19]. The sets of the players' pure strategies are

$$X = \{x(t), t \in [t_1; t_2],$$

$$t_1 < t_2 : a_{\min} \leq x(t) \leq a_{\max} \text{ by } a_{\min} < a_{\max}\} \subset \mathbb{L}_2[t_1; t_2] \quad (3)$$

and

$$Y = \{y(t), t \in [t_1; t_2],$$

$$t_1 < t_2 : b_{\min} \leq y(t) \leq b_{\max} \text{ by } b_{\min} < b_{\max}\} \subset \mathbb{L}_2[t_1; t_2], \quad (4)$$

respectively. Each of sets (3) and (4) is a rectangular functional space, in which every element is a bounded function of time by (1) and (2).

The first player's payoff in situation

$$\{x(t), y(t)\} \quad (5)$$

is

$$K(x(t), y(t)).$$

The payoff is presumed to be an integral functional:

$$K(x(t), y(t)) = \int_{[t_1; t_2]} f(x(t), y(t), t) d\mu(t) \quad (6)$$

with a function

$$f(x(t), y(t), t) \quad (7)$$

of $x(t)$ and $y(t)$ explicitly including time t . Therefore, the continuous zero-sum game

$$\langle \{X, Y\}, K(x(t), y(t)) \rangle \quad (8)$$

is defined on product

$$X \times Y \subset \mathbb{L}_2[t_1; t_2] \times \mathbb{L}_2[t_1; t_2] \quad (9)$$

of rectangular functional spaces (3) and (4) of players' pure strategies. It is worth noting that the game continuity is defined by the continuity of spaces (3) and (4), whereas payoff functional (6) still can have discontinuities.

As it has been argued above, zero-sum game (8), in which the players' strategies are functions of time, in practical reality is played discretely during time interval $[t_1, t_2]$. The time step is the same for each of the players because it is presumed to be established either by the rules of the system game-modelled or by the administrator.

A zero-sum game with staircase-function strategies

Denote by N the number of subintervals at which the player's pure strategy is constant, where $N \in \mathbb{N} \setminus \{1\}$. Then the player's pure strategy is a staircase function having only N different values (out of, maybe, a continuum of possible values). Then there are $N - 1$ time points at which the staircase-function strategy changes or can change its value. These points are $\{\tau^{(i)}\}_{i=1}^{N-1}$, where

$$t_1 = \tau^{(0)} < \tau^{(1)} < \tau^{(2)} < \dots < \tau^{(N-1)} < \tau^{(N)} = t_2. \quad (10)$$

Points $\{\tau^{(i)}\}_{i=0}^N$ are not necessarily to be equidistant, but they are the same for each of the players and they do not change as the game is repeated (a finite number of repetitions is meant – from the practical point of view).

The staircase-function strategies are right-continuous:

$$\lim_{\substack{\varepsilon > 0 \\ \varepsilon \rightarrow 0}} x(\tau^{(i)} + \varepsilon) = x(\tau^{(i)}) \quad (11)$$

and

$$\lim_{\substack{\varepsilon > 0 \\ \varepsilon \rightarrow 0}} y(\tau^{(i)} + \varepsilon) = y(\tau^{(i)}) \quad (12)$$

for $i = \overline{1, N-1}$, whereas (if the strategy value changes)

$$\lim_{\substack{\varepsilon > 0 \\ \varepsilon \rightarrow 0}} x(\tau^{(i)} - \varepsilon) \neq x(\tau^{(i)}) \quad (13)$$

and

$$\lim_{\substack{\varepsilon > 0 \\ \varepsilon \rightarrow 0}} y(\tau^{(i)} - \varepsilon) \neq y(\tau^{(i)}) \quad (14)$$

for $i = \overline{1, N-1}$. As an exception,

$$\lim_{\substack{\varepsilon > 0 \\ \varepsilon \rightarrow 0}} x(\tau^{(N)} - \varepsilon) = x(\tau^{(N)}) \quad (15)$$

and

$$\lim_{\substack{\varepsilon > 0 \\ \varepsilon \rightarrow 0}} y(\tau^{(N)} - \varepsilon) = y(\tau^{(N)}), \quad (16)$$

so

$$x(\tau^{(N-1)}) = x(\tau^{(N)}) \quad (17)$$

and

$$y(\tau^{(N-1)}) = y(\tau^{(N)}). \quad (18)$$

As both functions $x(t)$ and $y(t)$ are constant

$$\forall t \in [\tau^{(i-1)}; \tau^{(i)}] \text{ for } i = \overline{1, N-1} \text{ and} \\ \forall t \in [\tau^{(N-1)}; \tau^{(N)}],$$

then game (8) can be thought of as it is a succession of N continuous zero-sum games

$$\langle \{[a_{\min}; a_{\max}], [b_{\min}; b_{\max}]\}, K(\alpha_i, \beta_i) \rangle \quad (19)$$

defined on product

$$[a_{\min}; a_{\max}] \times [b_{\min}; b_{\max}]$$

by

$$\alpha_i = x(t) \in [a_{\min}; a_{\max}]$$

$$\text{and } \beta_i = y(t) \in [b_{\min}; b_{\max}] \quad \forall t \in [\tau^{(i-1)}; \tau^{(i)})$$

$$\text{for } i = \overline{1, N-1} \text{ and } \forall t \in [\tau^{(N-1)}; \tau^{(N)}] \quad (20)$$

where the factual payoff in situation

$$\{\alpha_i, \beta_i\} \quad (21)$$

is

$$K(\alpha_i, \beta_i) = \int_{[\tau^{(i-1)}; \tau^{(i)})} f(\alpha_i, \beta_i, t) d\mu(t) \\ \forall i = \overline{1, N-1} \quad (22)$$

and

$$K(\alpha_N, \beta_N) = \int_{[\tau^{(N-1)}; \tau^{(N)}]} f(\alpha_N, \beta_N, t) d\mu(t). \quad (23)$$

Henceforward, game (8) equivalent to the succession of N continuous zero-sum games (19) by (20)–(23) is called staircase. A pure-strategy situation in staircase game (8) is a succession of N situations

$$\{\{\alpha_i, \beta_i\}\}_{i=1}^N \quad (24)$$

in games (19). In staircase game (8), the set of the player's pure strategies is still a continuum of staircase functions of time, but the time is discrete. This time-discretization property, implying constant values of the players' strategies on every subinterval, allows, in addition to the succession of N continuous zero-sum games (19), decomposing staircase game (8) with respect to the (staircase) payoff.

Theorem 1. In a pure-strategy situation (5) of staircase game (8), represented as a succession of N games (19), functional (6) is re-written as a subinterval-wise sum

$$K(x(t), y(t)) = \sum_{i=1}^N K(\alpha_i, \beta_i) = \\ = \sum_{i=1}^{N-1} \int_{[\tau^{(i-1)}; \tau^{(i)})} f(\alpha_i, \beta_i, t) d\mu(t) + \\ + \int_{[\tau^{(N-1)}; \tau^{(N)}]} f(\alpha_N, \beta_N, t) d\mu(t). \quad (25)$$

Proof. Situation (21) is tied to half-subinterval

$$[\tau^{(i-1)}; \tau^{(i)}) \text{ by } i = \overline{1, N-1}$$

and to subinterval

$$[\tau^{(N-1)}; \tau^{(N)}] \text{ by } i = N.$$

Function (7) in this situation is some function of time t . Denote this function by $\psi_i(t)$. For situation (21) function

$$\psi_i(t) = 0 \quad \forall t \notin [\tau^{(i-1)}; \tau^{(i)}), \quad (26)$$

and for situation

$$\{\alpha_N, \beta_N\}$$

function

$$\psi_N(t) = 0 \quad \forall t \notin [\tau^{(N-1)}; \tau^{(N)}]. \quad (27)$$

Therefore,

$$f(x(t), y(t), t) = \sum_{i=1}^N \psi_i(t) \quad (28)$$

in a pure-strategy situation (5) of staircase game (8), by using (26) and (27). Consequently,

$$\begin{aligned} K(x(t), y(t)) &= \int_{[t_1; t_2]} f(x(t), y(t), t) d\mu(t) = \\ &= \sum_{i=1}^{N-1} \int_{[\tau^{(i-1)}; \tau^{(i)}]} \psi_i(t) d\mu(t) + \int_{[\tau^{(N-1)}; \tau^{(N)}]} \psi_N(t) d\mu(t) = \\ &= \sum_{i=1}^{N-1} \int_{[\tau^{(i-1)}; \tau^{(i)}]} f(\alpha_i, \beta_i, t) d\mu(t) + \\ &+ \int_{[\tau^{(N-1)}; \tau^{(N)}]} f(\alpha_N, \beta_N, t) d\mu(t) = \\ &= \sum_{i=1}^N K(\alpha_i, \beta_i) \end{aligned} \quad (29)$$

in a pure-strategy situation (5) of staircase game (8).

It is noteworthy that Theorem 1 can be proved also by using the property of countable additivity of the Lebesgue integral. Theorem 1 does not provide a method of solving the staircase game. Nevertheless, it provides a fundamental decomposition of the game based on the subinterval-wise summing in (25). This decomposition allows considering and solving each game (19) separately, whereupon the solutions are stitched (stacked) together.

Why the sampling must be different and irregular

There are two main arguments for considering different sampling steps at each of the players. First, the players cannot agree on the sampling step. If even they have identical ranges of function-strategy values, an agreement is impossible due to the cooperation is excluded. Second, if a player has a wider range of one's function-strategy values then it is likely to be sampled with a greater number of points. This, however, does not mean a denser sampling.

The sampling can be non-uniform (irregular). Indeed, a player may tend to use greater or lesser values of one's function-strategy more frequently. In particular, this may lead to a denser sampling in a neighbourhood of those values. Thus, the sampling, in a generalized approach to finite approximation of zero-sum games played in staircase-function continuous spaces, must be irregular.

Sampling along the pure strategy value axis

In game (19) on subinterval i , the first player has its set $[a_{\min}; a_{\max}]$ of pure strategies, and the second player's pure strategy set is $[b_{\min}; b_{\max}]$. Let set $[a_{\min}; a_{\max}]$ be sampled non-uniformly (irregularly) with M points, $M \in \mathbb{N} \setminus \{1\}$:

$$\begin{aligned} A(M) &= \{a^{(m)}\}_{m=1}^M = \\ &= \{a_{\min}, \{a^{(m)}\}_{m=2}^{M-1}, a_{\max}\} \subset [a_{\min}; a_{\max}] \end{aligned} \quad (30)$$

by

$$a^{(1)} = a_{\min} \quad \text{and} \quad a^{(M)} = a_{\max}, \quad (31)$$

i. e., the endpoints are always included into the sampling. Similarly to this, let set $[b_{\min}; b_{\max}]$ be sampled non-uniformly (irregularly) with J points, $J \in \mathbb{N} \setminus \{1\}$:

$$\begin{aligned} B(J) &= \{b^{(j)}\}_{j=1}^J = \\ &= \{b_{\min}, \{b^{(j)}\}_{j=2}^{J-1}, b_{\max}\} \subset [b_{\min}; b_{\max}] \end{aligned} \quad (32)$$

by

$$b^{(1)} = b_{\min} \quad \text{and} \quad b^{(J)} = b_{\max}. \quad (33)$$

The roughest sampling is by $M = 2$ and $J = 2$, when

$$A(2) = \{a^{(1)}, a^{(2)}\} = \{a_{\min}, a_{\max}\}$$

and

$$B(2) = \{b^{(1)}, b^{(2)}\} = \{b_{\min}, b_{\max}\}.$$

If either of integers M and J is increased by 1, a new sampling must comply with the previous one. This is a requirement of the proper sampling increment.

Definition 1. Sampling

$$\begin{aligned} \Psi(S+1) &= \{\lambda^{(s)}\}_{s=1}^{S+1} = \\ &= \{\zeta_{\min}, \{\lambda^{(s)}\}_{s=2}^S, \zeta_{\max}\} \subset [\zeta_{\min}; \zeta_{\max}] \end{aligned} \quad (34)$$

by $\zeta_{\min} < \zeta_{\max}$ and $S \in \mathbb{N} \setminus \{1\}$ is a proper sampling increment of sampling

$$\begin{aligned} \Psi(S) &= \{\zeta^{(s)}\}_{s=1}^S = \\ &= \{\zeta_{\min}, \{\zeta^{(s)}\}_{s=2}^{S-1}, \zeta_{\max}\} \subset [\zeta_{\min}; \zeta_{\max}] \end{aligned} \quad (35)$$

if

$$\max_{s=1, S} (\lambda^{(s+1)} - \lambda^{(s)}) < \max_{s=1, S-1} (\zeta^{(s+1)} - \zeta^{(s)}), \quad (36)$$

i. e. the $S + 1$ points in 1-incremented sampling (34) are selected denser than S points in sampling (35).

With the sampling by (30)–(33), the succession of N continuous games (19) by (20), (22), (23) becomes a succession of N matrix $M \times J$ games

$$\left\langle \left\{ \left\{ a^{(m)} \right\}_{m=1}^M, \left\{ b^{(j)} \right\}_{j=1}^J \right\}, \mathbf{K}_i(M, J) \right\rangle \quad (37)$$

with payoff matrices

$$\mathbf{K}_i(M, J) = [k_{imj}(M, J)]_{M \times J} \quad (38) \text{ and}$$

whose elements are

$$k_{imj}(M, J) = \int_{[\tau^{(i-1)}, \tau^{(i)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) \quad (39)$$

for $i = \overline{1, N-1}$

and

$$k_{Nmj}(M, J) = \int_{[\tau^{(N-1)}, \tau^{(N)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t). \quad (40)$$

So, if integers M and J for game (8) by (20) is somehow selected, the staircase game is represented as a succession of N matrix $M \times J$ games (37).

By sampling (30) and (32) the staircase game becomes defined on product $A(M) \times B(J)$, which becomes a product of staircase-function finite spaces by running through all $i = \overline{1, N}$. Thus, game (8) becomes a finite staircase game. It might be rendered to a matrix game in order to obtain a staircase solution (herein, adjective “staircase” gives a hint to the type of the game, rather than to the structure of its solution). However, there is a much easier way to solve a finite staircase game.

Theorem 2. If game (8) on product (9) by conditions (3), (4), (6) is made a staircase game as a succession of N continuous zero-sum games (19) by (20), (22), (23), whereupon it is sampled by (30) and (32), then the respective finite staircase game is always solved as a stack of successive optimal solutions of N matrix games (37) by (38)–(40).

Proof. A matrix game always has a solution, either in pure or mixed strategies. Denote by

$$\mathbf{P}_i(M, J) = [p_i^{(m)}(M, J)]_{1 \times M}$$

and

$$\mathbf{Q}_i(M, J) = [q_i^{(j)}(M, J)]_{1 \times J}$$

the mixed strategies of the first and second players, respectively, in matrix game (37). The respective sets of mixed strategies of the first and second players are

$$\mathbf{P} = \left\{ \mathbf{P}_i(M, J) \in \mathbb{R}^M : p_i^{(m)}(M, J) \geq 0, \sum_{m=1}^M p_i^{(m)}(M, J) = 1 \right\} \quad (41)$$

and

$$\mathbf{Q} = \left\{ \mathbf{Q}_i(M, J) \in \mathbb{R}^J : q_i^{(j)}(M, J) \geq 0, \sum_{j=1}^J q_i^{(j)}(M, J) = 1 \right\}, \quad (42)$$

so

$$\mathbf{P}_i(M, J) \in \mathbf{P}, \quad \mathbf{Q}_i(M, J) \in \mathbf{Q},$$

and

$$\{\mathbf{P}_i(M, J), \mathbf{Q}_i(M, J)\} \quad (43)$$

is a situation in this game, i. e. (43) is a situation on subinterval i . Let

$$\begin{aligned} & \left\{ \left\{ \mathbf{P}_i^*(M, J), \mathbf{Q}_i^*(M, J) \right\}_{i=1}^N \right\} = \\ & = \left\{ \left\{ [p_i^{(m)*}(M, J)]_{1 \times M}, [q_i^{(j)*}(M, J)]_{1 \times J} \right\}_{i=1}^N \right\} \end{aligned} \quad (44)$$

be a set of optimal solutions of N games (37) by (39) and (40). Then

$$\begin{aligned} & \max_{\mathbf{P}_i(M, J) \in \mathbf{P}} \min_{\mathbf{Q}_i(M, J) \in \mathbf{Q}} \mathbf{P}_i(M, J) \cdot \mathbf{K}_i(M, J) \cdot [\mathbf{Q}_i(M, J)]^T = \\ & = \max_{\mathbf{P}_i(M, J) \in \mathbf{P}} \min_{\mathbf{Q}_i(M, J) \in \mathbf{Q}} \sum_{m=1}^M \sum_{j=1}^J k_{imj}(M, J) \times \\ & \quad \times p_i^{(m)}(M, J) q_i^{(j)}(M, J) = \\ & = \max_{\mathbf{P}_i(M, J) \in \mathbf{P}} \min_{\mathbf{Q}_i(M, J) \in \mathbf{Q}} \sum_{m=1}^M \sum_{j=1}^J p_i^{(m)}(M, J) q_i^{(j)}(M, J) \times \\ & \quad \times \int_{[\tau^{(i-1)}, \tau^{(i)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) = \end{aligned}$$

$$\begin{aligned}
&= \sum_{m=1}^M \sum_{j=1}^J p_i^{(m)*}(M, J) q_i^{(j)*}(M, J) \times \\
&\quad \times \int_{[\tau^{(i-1)}; \tau^{(i)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) = \\
&= \mathbf{P}_i^*(M, J) \cdot \mathbf{K}_i(M, J) \cdot [\mathbf{Q}_i^*(M, J)]^T = v_i^*(M, J) = \\
&= \min_{\mathbf{Q}_i(M, J) \in \mathcal{Q}} \max_{\mathbf{P}_i(M, J) \in \mathcal{P}} \sum_{m=1}^M \sum_{j=1}^J p_i^{(m)}(M, J) q_i^{(j)}(M, J) \times \\
&\quad \times \int_{[\tau^{(i-1)}; \tau^{(i)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) = \\
&= \min_{\mathbf{Q}_i(M, J) \in \mathcal{Q}} \max_{\mathbf{P}_i(M, J) \in \mathcal{P}} \sum_{m=1}^M \sum_{j=1}^J k_{mj}(M, J) \times \\
&\quad \times p_i^{(m)}(M, J) q_i^{(j)}(M, J) = \\
&= \min_{\mathbf{Q}_i(M, J) \in \mathcal{Q}} \max_{\mathbf{P}_i(M, J) \in \mathcal{P}} \mathbf{P}_i(M, J) \cdot \mathbf{K}_i(M, J) \cdot [\mathbf{Q}_i(M, J)]^T \\
&\quad \forall i = \overline{1, N-1} \quad (45)
\end{aligned}$$

and

$$\begin{aligned}
&\max_{\mathbf{P}_N(M, J) \in \mathcal{P}} \min_{\mathbf{Q}_N(M, J) \in \mathcal{Q}} \mathbf{P}_N(M, J) \cdot \mathbf{K}_N(M, J) \cdot [\mathbf{Q}_N(M, J)]^T = \\
&= \max_{\mathbf{P}_N(M, J) \in \mathcal{P}} \min_{\mathbf{Q}_N(M, J) \in \mathcal{Q}} \sum_{m=1}^M \sum_{j=1}^J k_{Nmj}(M, J) \times \\
&\quad \times p_N^{(m)}(M, J) q_N^{(j)}(M, J) = \\
&= \max_{\mathbf{P}_N(M, J) \in \mathcal{P}} \min_{\mathbf{Q}_N(M, J) \in \mathcal{Q}} \sum_{m=1}^M \sum_{j=1}^J p_N^{(m)}(M, J) q_N^{(j)}(M, J) \times \\
&\quad \times \int_{[\tau^{(N-1)}; \tau^{(N)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) = \\
&= \sum_{m=1}^M \sum_{j=1}^J p_N^{(m)*}(M, J) q_N^{(j)*}(M, J) \times \\
&\quad \times \int_{[\tau^{(N-1)}; \tau^{(N)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) = \\
&= \mathbf{P}_N^*(M, J) \cdot \mathbf{K}_N(M, J) \cdot [\mathbf{Q}_N^*(M, J)]^T = v_N^*(M, J) = \\
&= \min_{\mathbf{Q}_N(M, J) \in \mathcal{Q}} \max_{\mathbf{P}_N(M, J) \in \mathcal{P}} \sum_{m=1}^M \sum_{j=1}^J p_N^{(m)}(M, J) q_N^{(j)}(M, J) \times \\
&\quad \times \int_{[\tau^{(N-1)}; \tau^{(N)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) =
\end{aligned}$$

$$\begin{aligned}
&= \min_{\mathbf{Q}_N(M, J) \in \mathcal{Q}} \max_{\mathbf{P}_N(M, J) \in \mathcal{P}} \sum_{m=1}^M \sum_{j=1}^J k_{Nmj}(M, J) \times \\
&\quad \times p_N^{(m)}(M, J) q_N^{(j)}(M, J) = \\
&= \min_{\mathbf{Q}_N(M, J) \in \mathcal{Q}} \max_{\mathbf{P}_N(M, J) \in \mathcal{P}} \mathbf{P}_N(M, J) \times \\
&\quad \times \mathbf{K}_N(M, J) \cdot [\mathbf{Q}_N(M, J)]^T. \quad (46)
\end{aligned}$$

Using Theorem 1 allows concluding that, using a wide-sense pure-strategy symbolism,

$$\begin{aligned}
&\max_{x(t) \in X} \min_{y(t) \in Y} K(x(t), y(t)) = \\
&= \sum_{i=1}^{N-1} \left(\max_{\mathbf{P}_i(M, J) \in \mathcal{P}} \min_{\mathbf{Q}_i(M, J) \in \mathcal{Q}} \sum_{m=1}^M \sum_{j=1}^J p_i^{(m)}(M, J) \times \right. \\
&\quad \times q_i^{(j)}(M, J) \int_{[\tau^{(i-1)}; \tau^{(i)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) \left. + \right. \\
&\quad \left. + \max_{\mathbf{P}_N(M, J) \in \mathcal{P}} \min_{\mathbf{Q}_N(M, J) \in \mathcal{Q}} \sum_{m=1}^M \sum_{j=1}^J p_N^{(m)}(M, J) \times \right. \\
&\quad \times q_N^{(j)}(M, J) \int_{[\tau^{(N-1)}; \tau^{(N)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) = \\
&= \sum_{i=1}^{N-1} \sum_{m=1}^M \sum_{j=1}^J p_i^{(m)*}(M, J) q_i^{(j)*}(M, J) \times \\
&\quad \times \int_{[\tau^{(i-1)}; \tau^{(i)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) + \\
&\quad + \sum_{m=1}^M \sum_{j=1}^J p_N^{(m)*}(M, J) q_N^{(j)*}(M, J) \times \\
&\quad \times \int_{[\tau^{(N-1)}; \tau^{(N)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) = \\
&= \sum_{i=1}^{N-1} \mathbf{P}_i^*(M, J) \cdot \mathbf{K}_i(M, J) \cdot [\mathbf{Q}_i^*(M, J)]^T + \\
&\quad + \mathbf{P}_N^*(M, J) \cdot \mathbf{K}_N(M, J) \cdot [\mathbf{Q}_N^*(M, J)]^T = \\
&= \sum_{i=1}^{N-1} v_i^*(M, J) + v_N^*(M, J) = v^*(M, J) =
\end{aligned}$$

$$\begin{aligned}
 &= \sum_{i=1}^{N-1} \left(\min_{\mathbf{Q}_i(M, J) \in \mathcal{Q}} \max_{\mathbf{P}_i(M, J) \in \mathcal{P}} \sum_{m=1}^M \sum_{j=1}^J p_i^{(m)}(M, J) \times \right. \\
 &\quad \times q_i^{(j)}(M, J) \int_{[\tau^{(i-1)}, \tau^{(i)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) \left. + \right. \\
 &\quad + \min_{\mathbf{Q}_N(M, J) \in \mathcal{Q}} \max_{\mathbf{P}_N(M, J) \in \mathcal{P}} \sum_{m=1}^M \sum_{j=1}^J p_N^{(m)}(M, J) \times \\
 &\quad \times q_N^{(j)}(M, J) \int_{[\tau^{(N-1)}, \tau^{(N)}]} f(a^{(m)}, b^{(j)}, t) d\mu(t) = \\
 &\quad = \min_{y(t) \in Y} \max_{x(t) \in X} K(x(t), y(t)) \quad (47)
 \end{aligned}$$

and, therefore, the stack of successive solutions (44) is an optimal solution in staircase game (8) sampled by (30) and (32).

It is quite clear that these $M \times J$ matrix games are solved in parallel, without caring of the succession. The succession does matter when the solutions are stacked (stitched) together to form the staircase solution (the solution to the finite staircase game). If all N matrix games are solved in pure strategies, then stacking their solutions is fulfilled trivially. When there is at least an equilibrium in mixed strategies for a subinterval, the stacking is fulfilled as well implying that the resulting pure-mixed-strategy solution of staircase game (8) is realized successively, subinterval by subinterval, spending the same amount of time to implement both pure strategy and mixed strategy solutions (e. g., see [6, 7, 10, 18, 20, 21]).

Approximate solution consistency

The conditions of the appropriate finite approximation are stated by using the known method of obtaining the approximate solution of continuous antagonistic games on unit multidimensional cube with uniform sampling [18]. There are five items of the conditions. The requirement of the smooth sampling of the payoff kernel is inapplicable here [22].

First of all, there is an easy-to-find condition of the finite approximation appropriateness. It is about the game optimal value change, which must not change more by the proper sampling increment. Inasmuch as an increment is possible from the side of both the players, then this condition is a set of $3N$ inequalities:

$$\begin{aligned}
 |v_i^*(M, J) - v_i^*(M + 1, J)| &\leq |v_i^*(M - 1, J) - v_i^*(M, J)| \\
 &\text{for } i = \overline{1, N} \quad (48)
 \end{aligned}$$

and

$$\begin{aligned}
 |v_i^*(M, J) - v_i^*(M, J + 1)| &\leq |v_i^*(M, J - 1) - v_i^*(M, J)| \\
 &\text{for } i = \overline{1, N} \quad (49)
 \end{aligned}$$

and

$$\begin{aligned}
 |v_i^*(M, J) - v_i^*(M + 1, J + 1)| &\leq |v_i^*(M - 1, J - 1) - v_i^*(M, J)| \\
 &\text{for } i = \overline{1, N}. \quad (50)
 \end{aligned}$$

Conditions (48)–(50) mean that, as the sampling density minimally increases, either from the side of the first or second player, the game optimal value change in an appropriate approximation should not grow.

Definition 2. An approximate solution (44) to staircase game (8) is called payoff- $\{M, J\}$ -consistent if inequalities (48)–(50) hold. The players’ optimal strategies in such a solution are called payoff- $\{M, J\}$ -consistent.

The second condition is the change of the optimal strategy support cardinality. Denote the supports of the optimal strategies of the players by

$$\text{supp } \mathbf{P}_i^*(M, J) = \{m_u\}_{u=1}^{U_i(M, J)} \subset \{m\}_{m=1}^M \quad (51)$$

by the respective support probabilities

$$\{p_i^{(m_u)^*}(M, J)\}_{u=1}^{U_i(M, J)} \quad (52)$$

and

$$\text{supp } \mathbf{Q}_i^*(M, J) = \{j_w\}_{w=1}^{W_i(M, J)} \subset \{j\}_{j=1}^J \quad (53)$$

by the respective support probabilities

$$\{q_i^{(j_w)^*}(M, J)\}_{w=1}^{W_i(M, J)}. \quad (54)$$

Then $6N$ inequalities

$$U_i(M + 1, J) \geq U_i(M, J) \quad \text{for } i = \overline{1, N}, \quad (55)$$

$$U_i(M, J + 1) \geq U_i(M, J) \quad \text{for } i = \overline{1, N}, \quad (56)$$

$$U_i(M + 1, J + 1) \geq U_i(M, J) \quad \text{for } i = \overline{1, N}, \quad (57)$$

$$W_i(M + 1, J) \geq W_i(M, J) \quad \text{for } i = \overline{1, N}, \quad (58)$$

$$W_i(M, J + 1) \geq W_i(M, J) \quad \text{for } i = \overline{1, N}, \quad (59)$$

$$W_i(M + 1, J + 1) \geq W_i(M, J) \quad \text{for } i = \overline{1, N} \quad (60)$$

require that, by minimally increasing the sampling density, either from the side of the first or second player, the cardinalities of the supports not decrease.

Definition 3. An approximate solution (44) to staircase game (8) is called weakly support-cardinality- $\{M, J\}$ -consistent if inequalities (55)–(60) hold. The players' optimal strategies in such a solution are called weakly support-cardinality- $\{M, J\}$ -consistent.

Obviously, requirements (55)–(60) can be supplemented (strengthened) by considering a minimal decrement of the sampling density. Then another $6N$ inequalities

$$U_i(M, J) \geq U_i(M-1, J) \quad \text{for } i = \overline{1, N}, \quad (61)$$

$$U_i(M, J) \geq U_i(M, J-1) \quad \text{for } i = \overline{1, N}, \quad (62)$$

$$U_i(M, J) \geq U_i(M-1, J-1) \quad \text{for } i = \overline{1, N}, \quad (63)$$

$$W_i(M, J) \geq W_i(M-1, J) \quad \text{for } i = \overline{1, N}, \quad (64)$$

$$W_i(M, J) \geq W_i(M, J-1) \quad \text{for } i = \overline{1, N}, \quad (65)$$

$$W_i(M, J) \geq W_i(M-1, J-1) \quad \text{for } i = \overline{1, N} \quad (66)$$

are required.

Definition 4. An approximate solution (44) to staircase game (8) is called support-cardinality- $\{M, J\}$ -consistent if inequalities (55)–(66) hold. The players' optimal strategies in such a solution are called support-cardinality- $\{M, J\}$ -consistent.

As the sampling density minimally increases, the maximal gap between the support indices should not increase. Let $m_u(M, J)$ and $j_w(M, J)$ be the respective support indices corresponding to integers $\{M, J\}$ on a subinterval by (20). Then $6N$ inequalities

$$\begin{aligned} & \frac{\max}{u=1, U_i(M+1, J)-1} [m_{u+1}(M+1, J) - m_u(M+1, J)] \leq \\ & \leq \frac{\max}{u=1, U_i(M, J)-1} [m_{u+1}(M, J) - m_u(M, J)] \quad \text{for } i = \overline{1, N}, \quad (67) \end{aligned}$$

$$\begin{aligned} & \frac{\max}{u=1, U_i(M, J+1)-1} [m_{u+1}(M, J+1) - m_u(M, J+1)] \leq \\ & \leq \frac{\max}{u=1, U_i(M, J)-1} [m_{u+1}(M, J) - m_u(M, J)] \quad \text{for } i = \overline{1, N}, \quad (68) \end{aligned}$$

$$\begin{aligned} & \frac{\max}{u=1, U_i(M+1, J+1)-1} [m_{u+1}(M+1, J+1) - m_u(M+1, J+1)] \leq \\ & \leq \frac{\max}{u=1, U_i(M, J)-1} [m_{u+1}(M, J) - m_u(M, J)] \quad \text{for } i = \overline{1, N}, \quad (69) \end{aligned}$$

$$\frac{\max}{w=1, W_i(M+1, J)-1} [j_{w+1}(M+1, J) - j_w(M+1, J)] \leq$$

$$\leq \frac{\max}{w=1, W_i(M, J)-1} [j_{w+1}(M, J) - j_w(M, J)] \quad \text{for } i = \overline{1, N}, \quad (70)$$

$$\begin{aligned} & \frac{\max}{w=1, W_i(M, J+1)-1} [j_{w+1}(M, J+1) - j_w(M, J+1)] \leq \\ & \leq \frac{\max}{w=1, W_i(M, J)-1} [j_{w+1}(M, J) - j_w(M, J)] \quad \text{for } i = \overline{1, N}, \quad (71) \end{aligned}$$

$$\begin{aligned} & \frac{\max}{w=1, W_i(M+1, J+1)-1} [j_{w+1}(M+1, J+1) - j_w(M+1, J+1)] \leq \\ & \leq \frac{\max}{w=1, W_i(M, J)-1} [j_{w+1}(M, J) - j_w(M, J)] \quad \text{for } i = \overline{1, N} \quad (72) \end{aligned}$$

are required.

Definition 5. An approximate solution (44) to staircase game (8) is called weakly sampling-density- $\{M, J\}$ -consistent if inequalities (67)–(72) hold. The players' optimal strategies in such a solution are called weakly sampling-density- $\{M, J\}$ -consistent.

Similarly to strengthening the weak (by Definition 3) support cardinality to that by Definition 4, requirements (67)–(72) can be strengthened by considering a minimal decrement of the sampling density. Then another $6N$ inequalities

$$\begin{aligned} & \frac{\max}{u=1, U_i(M, J)-1} [m_{u+1}(M, J) - m_u(M, J)] \leq \\ & \frac{\max}{u=1, U_i(M, J)-1} [m_{u+1}(M, J) - m_u(M, J)] \quad \text{for } i = \overline{1, N}, \quad (73) \end{aligned}$$

$$\begin{aligned} & \frac{\max}{u=1, U_i(M, J)-1} [m_{u+1}(M, J) - m_u(M, J)] \leq \\ & \leq \frac{\max}{u=1, U_i(M, J-1)-1} [m_{u+1}(M, J-1) - m_u(M, J-1)] \\ & \quad \text{for } i = \overline{1, N}, \quad (74) \end{aligned}$$

$$\begin{aligned} & \frac{\max}{u=1, U_i(M, J)-1} [m_{u+1}(M, J) - m_u(M, J)] \leq \\ & \leq \frac{\max}{u=1, U_i(M-1, J-1)-1} [m_{u+1}(M-1, J-1) - m_u(M-1, J-1)] \\ & \quad \text{for } i = \overline{1, N}, \quad (75) \end{aligned}$$

$$\begin{aligned} & \frac{\max}{w=1, W_i(M, J)-1} [j_{w+1}(M, J) - j_w(M, J)] \leq \\ & \leq \frac{\max}{w=1, W_i(M-1, J)-1} [j_{w+1}(M-1, J) - j_w(M-1, J)] \\ & \quad \text{for } i = \overline{1, N}, \quad (76) \end{aligned}$$

$$\begin{aligned} & \frac{\max}{w=1, W_i(M, J)-1} [j_{w+1}(M, J) - j_w(M, J)] \leq \\ & \leq \frac{\max}{w=1, W_i(M, J-1)-1} [j_{w+1}(M, J-1) - j_w(M, J-1)] \\ & \quad \text{for } i = \overline{1, N}, \quad (77) \end{aligned}$$

$$\begin{aligned} & \max_{w=1, \overline{W_i(M, J)-1}} [j_{w+1}(M, J) - j_w(M, J)] \leq \\ \leq & \max_{w=1, \overline{W_i(M-1, J-1)-1}} [j_{w+1}(M-1, J-1) - j_w(M-1, J-1)] \\ & \text{for } i = \overline{1, N} \end{aligned} \tag{78}$$

are required.

Definition 6. An approximate solution (44) to staircase game (8) is called sampling-density- $\{M, J\}$ -consistent if inequalities (67)–(78) hold. The players’ optimal strategies in such a solution are called sampling-density- $\{M, J\}$ -consistent.

Denote by $h_1(i; m, M, J)$ a polyline whose vertices are probabilities

$$\{p_i^{(m)*}(M, J)\}_{m=1}^M,$$

and denote by $h_2(i; j, M, J)$ a polyline whose vertices are probabilities

$$\{q_i^{(j)*}(M, J)\}_{j=1}^J.$$

Then, by minimally increasing the sampling density, the “neighboring” polylines should not be farther from each other, i. e. inequalities

$$\begin{aligned} & \max_{[0; 1]} |h_1(i; m, M, J) - h_1(i; m, M + 1, J)| \leq \\ & \leq \max_{[0; 1]} |h_1(i; m, M - 1, J) - h_1(i; m, M, J)| \\ & \text{for } i = \overline{1, N}, \end{aligned} \tag{79} \text{ and}$$

$$\begin{aligned} & \max_{[0; 1]} |h_1(i; m, M, J) - h_1(i; m, M, J + 1)| \leq \\ & \leq \max_{[0; 1]} |h_1(i; m, M, J - 1) - h_1(i; m, M, J)| \\ & \text{for } i = \overline{1, N}, \end{aligned} \tag{80}$$

$$\begin{aligned} & \max_{[0; 1]} |h_1(i; m, M, J) - h_1(i; m, M + 1, J + 1)| \leq \\ & \leq \max_{[0; 1]} |h_1(i; m, M - 1, J - 1) - h_1(i; m, M, J)| \\ & \text{for } i = \overline{1, N}, \end{aligned} \tag{81}$$

and

$$\begin{aligned} & \max_{[0; 1]} |h_2(i; j, M, J) - h_2(i; j, M + 1, J)| \leq \\ & \leq \max_{[0; 1]} |h_2(i; j, M - 1, J) - h_2(i; j, M, J)| \\ & \text{for } i = \overline{1, N}, \end{aligned} \tag{82}$$

$$\begin{aligned} & \max_{[0; 1]} |h_2(i; j, M, J) - h_2(i; j, M, J + 1)| \leq \\ & \leq \max_{[0; 1]} |h_2(i; j, M, J - 1) - h_2(i; j, M, J)| \\ & \text{for } i = \overline{1, N}, \end{aligned} \tag{83}$$

$$\begin{aligned} & \max_{[0; 1]} |h_2(i; j, M, J) - h_2(i; j, M + 1, J + 1)| \leq \\ & \leq \max_{[0; 1]} |h_2(i; j, M - 1, J - 1) - h_2(i; j, M, J)| \\ & \text{for } i = \overline{1, N}, \end{aligned} \tag{84}$$

along with

$$\begin{aligned} & \|h_1(i; m, M, J) - h_1(i; m, M + 1, J)\| \leq \\ & \leq \|h_1(i; m, M - 1, J) - h_1(i; m, M, J)\| \\ & \text{in } \mathbb{L}_2[0; 1] \text{ for } i = \overline{1, N}, \end{aligned} \tag{85}$$

$$\begin{aligned} & \|h_1(i; m, M, J) - h_1(i; m, M, J + 1)\| \leq \\ & \leq \|h_1(i; m, M, J - 1) - h_1(i; m, M, J)\| \\ & \text{in } \mathbb{L}_2[0; 1] \text{ for } i = \overline{1, N}, \end{aligned} \tag{86}$$

$$\begin{aligned} & \|h_1(i; m, M, J) - h_1(i; m, M + 1, J + 1)\| \leq \\ & \leq \|h_1(i; m, M - 1, J - 1) - h_1(i; m, M, J)\| \\ & \text{in } \mathbb{L}_2[0; 1] \text{ for } i = \overline{1, N}, \end{aligned} \tag{87}$$

and

$$\begin{aligned} & \|h_2(i; j, M, J) - h_2(i; j, M + 1, J)\| \leq \\ & \leq \|h_2(i; j, M - 1, J) - h_2(i; j, M, J)\| \\ & \text{in } \mathbb{L}_2[0; 1] \text{ for } i = \overline{1, N}, \end{aligned} \tag{88}$$

$$\begin{aligned} & \|h_2(i; j, M, J) - h_2(i; j, M, J + 1)\| \leq \\ & \leq \|h_2(i; j, M, J - 1) - h_2(i; j, M, J)\| \\ & \text{in } \mathbb{L}_2[0; 1] \text{ for } i = \overline{1, N}, \end{aligned} \tag{89}$$

$$\begin{aligned} & \|h_2(i; j, M, J) - h_2(i; j, M + 1, J + 1)\| \leq \\ & \leq \|h_2(i; j, M - 1, J - 1) - h_2(i; j, M, J)\| \\ & \text{in } \mathbb{L}_2[0; 1] \text{ for } i = \overline{1, N}. \end{aligned} \tag{90}$$

are required.

Definition 7. An approximate solution (44) to staircase game (8) is called probability- $\{M, J\}$ -consistent if inequalities (79)–(90) hold. The players’ optimal strategies in such a solution are called probability- $\{M, J\}$ -consistent.

The solution consistency by each of Definitions 3–7 implies that both the players' optimal strategies are consistent as well. Nevertheless, it is not worth cancelling the player's optimal strategy consistency when for the other player the consistency conditions do not hold. Thus, a player's optimal strategy may be consistent while an optimal strategy of the other player is not consistent. For instance, if inequalities (55)–(57) hold, but inequalities (58)–(60) do not hold (at least one of those $3N$ inequalities is violated) and thus the second player's optimal strategy is not weakly support-cardinality- $\{M, J\}$ -consistent, the first player's optimal strategy is weakly support-cardinality- $\{M, J\}$ -consistent. If inequalities (48)–(50), (55)–(60), (67)–(72), (79)–(90) hold for some i , then matrix game (37), assigned to the sub-interval between $\tau^{(i-1)}$ and $\tau^{(i)}$, has a weakly consistent approximate solution to the corresponding continuous game (19) by (20), (22), (23). On this basis, the weak consistency of an approximate solution to a staircase game (8) is formulated.

Definition 8. The stack of successive solutions (44) is called a weakly $\{M, J\}$ -consistent approximate solution of game (8) on product (9) by conditions (3), (4), (6) if inequalities (48)–(50), (55)–(60), (67)–(72), (79)–(90) hold. The players' optimal strategies in such a solution are called weakly $\{M, J\}$ -consistent.

Once again, if, say, the second player's optimal strategy is not weakly $\{M, J\}$ -consistent (at least one of the respective inequalities in Definition 8 for the second player is violated), it does not mean that the first player's optimal strategy is not weakly $\{M, J\}$ -consistent also. If, in this example, inequalities (48)–(50), (55)–(57), (67)–(69), (79)–(81), (85)–(87) do hold, then the first player possesses a weakly $\{M, J\}$ -consistent optimal strategy, regardless of the second player's weak consistency. Similarly to strengthening Definitions 3 and 5, the weak consistency can be strengthened by adding the requirements with inequalities (61)–(66) and (73)–(78).

Definition 9. The stack of successive solutions (44) is called an $\{M, J\}$ -consistent approximate solution of game (8) on product (9) by conditions (3), (4), (6) if inequalities (48)–(50) and (55)–(90) hold. The players' optimal strategies in such a solution are called $\{M, J\}$ -consistent.

The approximate solution consistency theoretically proposes a better approximation than the weak consistency. The weak consistency notion by Definition 8 may be thought of as it is decomposed by Definitions 2, 3, 5, 7. Thus, the consistency notion by Definition 9 is decomposed into Definitions 2, 4, 6, 7. Even if an approximate solution is not weakly consistent, it may be, e. g., payoff-con-

sistent. The payoff consistency is checked the easiest and fastest. A payoff-consistent solution can be sufficient to accept it as an appropriate approximate solution [1, 6, 14, 18, 22]. However, if a one of $3N$ inequalities (48)–(50) is violated, even this type of consistency does not work. Meanwhile, the violation may be induced by a very small growth of the payoff change. Therefore, it is useful and practically reasonable to consider the payoff consistency adding a relaxation to inequalities (48)–(50).

Definition 10. An approximate solution (44) to staircase game (8) is called ε -payoff- $\{M, J\}$ -consistent if inequalities

$$\begin{aligned} & |v_i^*(M, J) - v_i^*(M + 1, J)| - \\ & - \varepsilon \leq |v_i^*(M - 1, J) - v_i^*(M, J)| \\ & \text{by some } \varepsilon > 0 \text{ for } i = \overline{1, N}. \end{aligned} \quad (91)$$

and

$$\begin{aligned} & |v_i^*(M, J) - v_i^*(M, J + 1)| - \\ & - \varepsilon \leq |v_i^*(M, J - 1) - v_i^*(M, J)| \\ & \text{by some } \varepsilon > 0 \text{ for } i = \overline{1, N}. \end{aligned} \quad (92)$$

and

$$\begin{aligned} & |v_i^*(M, J) - v_i^*(M + 1, J + 1)| \\ & - \varepsilon \leq |v_i^*(M - 1, J - 1) - v_i^*(M, J)| \\ & \text{by some } \varepsilon > 0 \text{ for } i = \overline{1, N}. \end{aligned} \quad (93)$$

hold. The players' optimal strategies in such a solution are called ε -payoff- $\{M, J\}$ -consistent.

To ascertain whether the stack of successive solutions (44) is weakly consistent or not, the seven bunches of N matrix games (37) should be solved, where the sampling density is defined by integers

$$\begin{aligned} & \{M - 1, J - 1\}, \{M - 1, J\}, \{M, J - 1\}, \{M, J\}, \\ & \{M + 1, J\}, \{M, J + 1\}, \{M + 1, J + 1\}. \end{aligned}$$

It is worth noting once again that the players select their respective integers M and J independently and, moreover, the sampling by an integer S means that those $S - 2$ points within an open interval can be chosen in any way, not necessarily to be uniformly distributed through the interval. Only the requirement of the proper sampling increment (by Definition 1) is followed. Nevertheless, the consistency meant by some sampling density integers $\{M, J\}$ does not guarantee that both the players will select such sampling density. Moreover, it is hard to find

a continuous zero-sum game, for which a consistent approximate solution could be determined at appropriately small integers M and J . However, it is quite naturally to expect that, as they are increased (i. e., the sampling is made denser), the approximate solutions must converge to the solution of staircase game (8). Besides, the approximate solutions must become “more” consistent, which means that more inequalities of the bunch of inequalities (48)–(50), (55)–(90) must hold.

A visual exemplification

To visually exemplify how a zero-sum staircase game is approximated by using the approximate solution consistency along with reconciling the difference of the players’ sampling step selection, consider a case in which $t \in [0.4\pi; 1.6\pi]$, the set of pure strategies of the first player is

$$X = \{x(t), t \in [0.4\pi; 1.6\pi] : 5 \leq x(t) \leq 9\} \subset \mathbb{L}_2 [0.4\pi; 1.6\pi] \tag{94}$$

and the set of pure strategies of the second player is

$$Y = \{y(t), t \in [0.4\pi; 1.6\pi] : 2 \leq y(t) \leq 8\} \subset \mathbb{L}_2 [0.4\pi; 1.6\pi], \tag{95}$$

where each of the players is allowed to change its pure strategy value at time points

$$\{\tau^{(i)}\}_{i=1}^5 = \{0.4\pi + 0.2\pi i\}_{i=1}^5.$$

The payoff functional is

$$K(x(t), y(t)) = \int_{[0.4\pi; 1.6\pi]} \sin^2\left(0.25xt + \frac{\pi}{11}\right) \times \sin\left(0.53yt - \frac{7\pi}{8}\right) e^{-0.02xt} d\mu(t). \tag{96}$$

So, each of the players possesses 6-subinterval staircase function-strategies defined on interval $[0.4\pi; 1.6\pi]$. Hence, the zero-sum staircase game is represented as a succession of 6 zero-sum games (19)

$$\langle \{[5; 9], [2; 8]\}, K(\alpha_i, \beta_i) \rangle \tag{97}$$

by

$$\alpha_i = x(t) \in [5; 9] \text{ and } \beta_i = y(t) \in [2; 8]$$

$$\forall t \in [0.2\pi + 0.2\pi i; 0.4\pi + 0.2\pi i] \text{ for } i = \overline{1, 5} \text{ and } \forall t \in [1.4\pi; 1.6\pi], \tag{98}$$

where the factual payoff in situation (21) is

$$K(\alpha_i, \beta_i) = \int_{[0.2\pi+0.2\pi i; 0.4\pi+0.2\pi i]} \sin^2\left(0.25\alpha_i t + \frac{\pi}{11}\right) \times \sin\left(0.53\beta_i t - \frac{7\pi}{8}\right) e^{-0.02\alpha_i t} d\mu(t) \quad \forall i = \overline{1, 5} \tag{99}$$

and

$$K(\alpha_6, \beta_6) = \int_{[1.4\pi; 1.6\pi]} \sin^2\left(0.25\alpha_6 t + \frac{\pi}{11}\right) \times \sin\left(0.53\beta_6 t - \frac{7\pi}{8}\right) e^{-0.02\alpha_6 t} d\mu(t). \tag{100}$$

Payoff functional (96) on each subinterval of set

$$\left\{ [0.2\pi + 0.2\pi i; 0.4\pi + 0.2\pi i] \right\}_{i=1}^5, [1.4\pi; 1.6\pi] \tag{101}$$

is shown in Fig. 1.

The irregularity (non-uniformity) in the sampling is modelled as follows:

$$a_0^{(m)} = 5 + \frac{4m-4}{M-1} \text{ and } a^{(m)} = a_0^{(m)} + \frac{\xi_1}{M} \text{ for } m = \overline{2, M-1} \tag{102}$$

by $a^{(1)} = 5$, $a^{(M)} = 9$, and

$$b_0^{(j)} = 2 + \frac{6j-6}{J-1} \text{ and } b^{(j)} = b_0^{(j)} + \frac{\xi_2}{J} \text{ for } j = \overline{2, J-1} \tag{103}$$

by $b^{(1)} = 2$, $b^{(J)} = 8$, where ξ_1 and ξ_2 are values of two independent random variables distributed normally with zero mean and unit variance. The values resulting from (102) and (103) are sorted in ascending order, whereupon they are checked whether (30) and (32) are true. When either integer M or J is increased by 1, samplings (30) and (32) are checked whether they satisfy the proper sampling increment by Definition 1, i. e. whether inequality (36) holds for samplings (35) and (34).

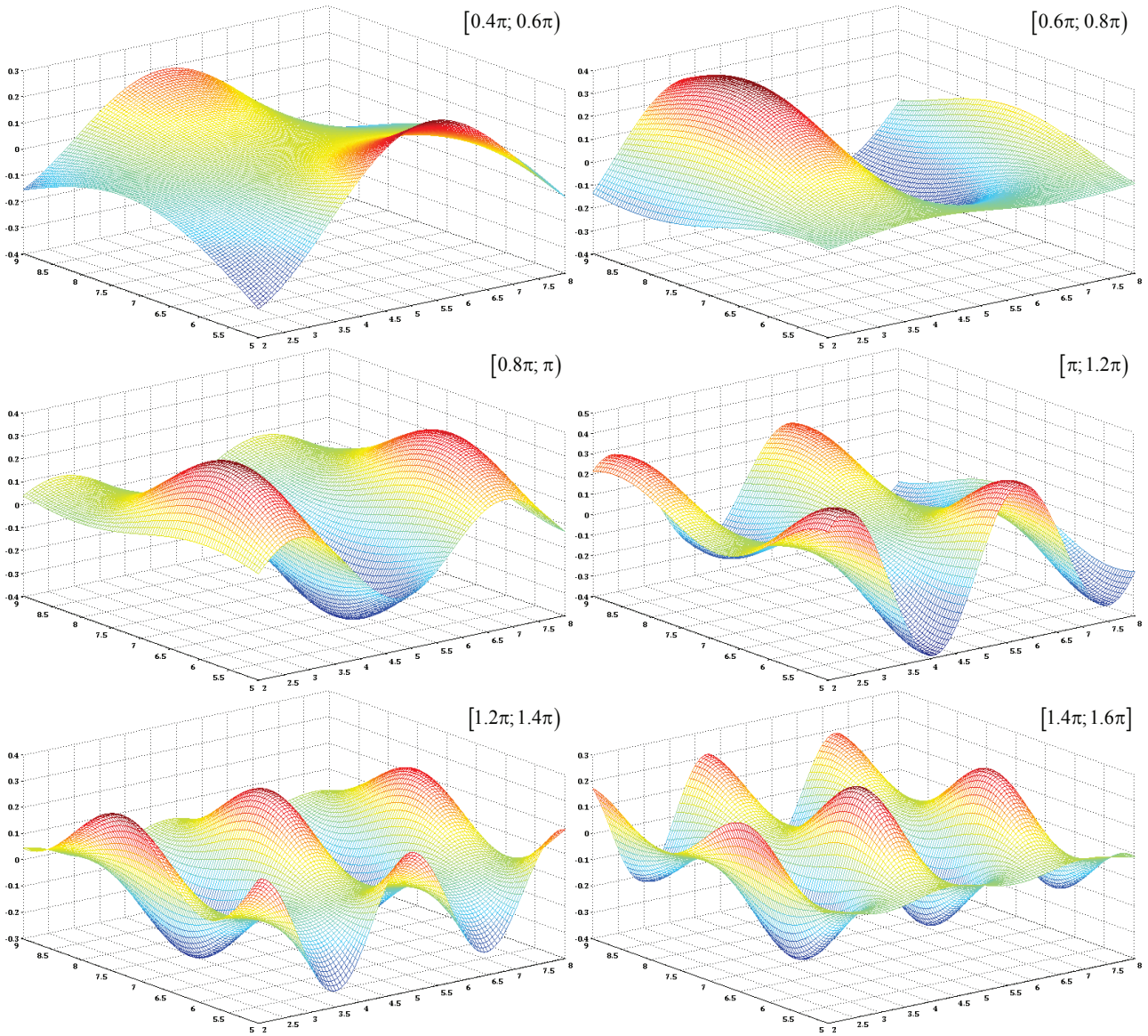


Fig. 1. The payoff kernels (99), (100) on the 6 subintervals of set (101)

Thus, 6 matrix games (37) with payoff matrices (38) are formed from 6 zero-sum games (97), where

$$k_{ij}(M, J) = \int_{[0.2\pi+0.2\pi i; 0.4\pi+0.2\pi i]} \sin^2\left(0.25a^{(m)}t + \frac{\pi}{11}\right) \times \sin\left(0.53b^{(j)}t - \frac{7\pi}{8}\right) e^{-0.02a^{(m)}t} d\mu(t) \text{ for } i = \overline{1, 5} \quad (104)$$

and

$$k_{6mj}(M, J) = \int_{[1.4\pi; 1.6\pi]} \sin^2\left(0.25a^{(m)}t + \frac{\pi}{11}\right) \times \sin\left(0.53b^{(j)}t - \frac{7\pi}{8}\right) e^{-0.02a^{(m)}t} d\mu(t). \quad (105)$$

Although the subinterval length in (104) and (105) does not change, every subinterval has its “own” matrix game due to time variable t is explicitly included into the functions under the integral. This means that, as time goes by, the players develop their actions subinterval by subinterval.

Surely, the solutions of these games (and the solution of the initial staircase game) badly depend on the sampling. Subinterval-wise optimal strategies of the players by the sampling for every $M = \overline{3, 10}$ and $J = \overline{3, 10}$ are shown in Fig. 2 in a bunch (they are presented indistinguishable). It is well seen that as the sampling density changes at such a relatively wide range of small sampling integers M and J , the player's optimal strategy (in every subinterval game, let alone the stacked optimal strategy on interval $[0.4\pi; 1.6\pi]$) badly varies. The only exception is the first subinterval, on which the second player's optimal strategy being the pure one does not vary at all. The first player has only pure optimal strategies on this subinterval as well. The first player's payoff $v_i^*(M, J)$ (at the end of the i -th subinterval) and the payoff cumulative sum

$$v^{(n)*}(M, J) = \sum_{i=1}^n v_i^*(M, J) \text{ by } n = \overline{1, 6} \quad (106)$$

are also badly scattered (Fig. 3), where, according to (106), $v^{(6)*}(M, J)$ is the optimal value in this staircase game, i. e.

$$v^*(M, J) = v^{(6)*}(M, J). \quad (107)$$

The only exception is that payoffs $v_6^*(M, J)$ received on reaching the final time point (at the end of the sixth subinterval) are almost converged (seen as dots), unlike optimal values (107) being scattered the worst (seen as circles).

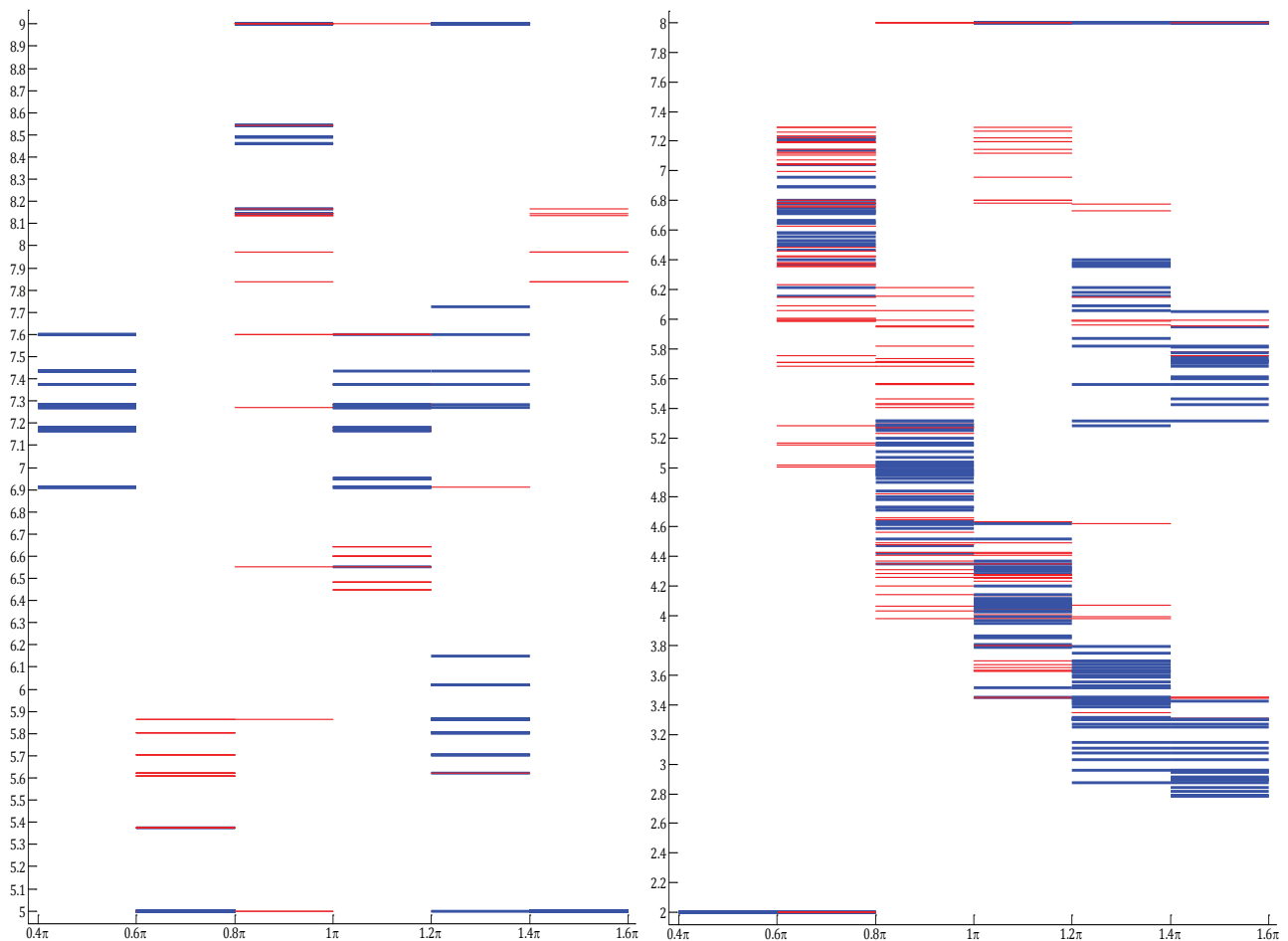


Fig. 2. An indistinguishable bunch of subinterval-wise optimal strategies of the first (left) and second (right) players by $M = \overline{3, 10}$ and $J = \overline{3, 10}$ (here and further below the optimal pure strategy is represented by thicker line, pure strategies from the mixed optimal strategy support are represented by thinner lines)

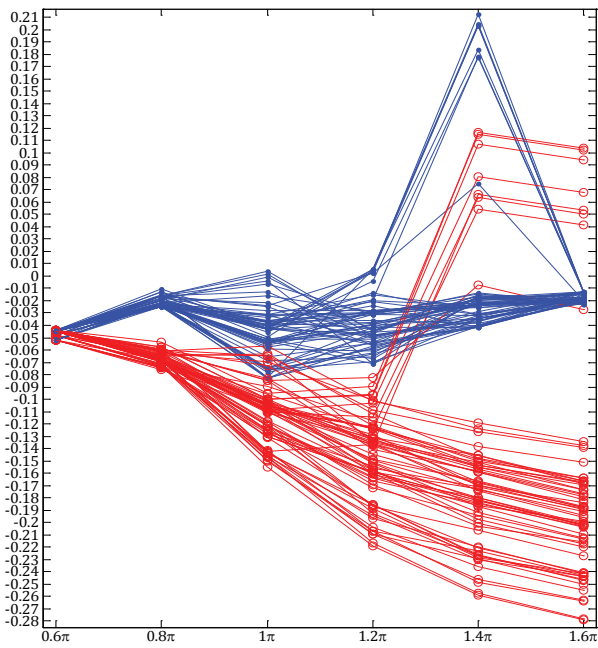


Fig. 3. An indistinguishable bunch of the first player's payoffs at the end of every subinterval (dots) and their cumulative sum (circles) by $M = \overline{3, 10}$ and $J = \overline{3, 10}$

Nevertheless, as the sampling density is further increased, mixed optimal strategies become more “condensed” (Fig. 4), where game (97) by (98)–(100) is still solved in pure strategies on the first subinterval. The pure optimal strategy of the second player on $[0.4\pi; 0.6\pi]$ is unchangeable (it is unchangeable, whichever the sampling is). Moreover, the payoffs “condense” also (Fig. 5): the subinterval payoffs run into a distinct polyline, and their cumulative sum runs into a polyline as well, although some scattering of optimal values (107) is still seen.

It is noteworthy that the players' optimal strategies are ε -payoff- $\{M, J\}$ -consistent just for

$$\varepsilon = 0.304 \cdot |v_i^*(M, J)| \quad \text{at } i = \overline{1, 6}$$

by every

$$M = \overline{15, 20} \quad \text{and} \quad J = \overline{15, 20}$$

This is an evidence of that the solution convergence is not enough.

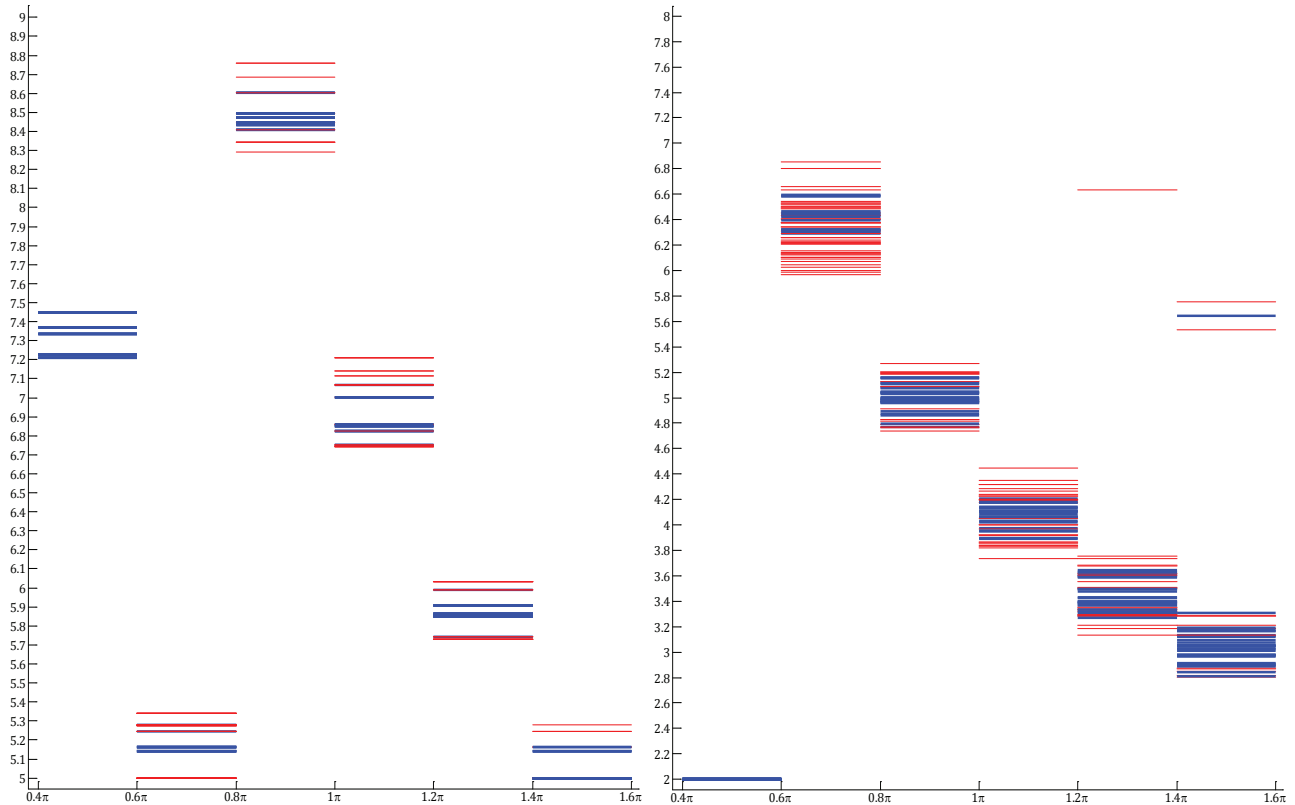


Fig. 4. A bunch of subinterval-wise optimal strategies of the first (left) and second (right) players by $M = \overline{15, 20}$ and $J = \overline{15, 20}$

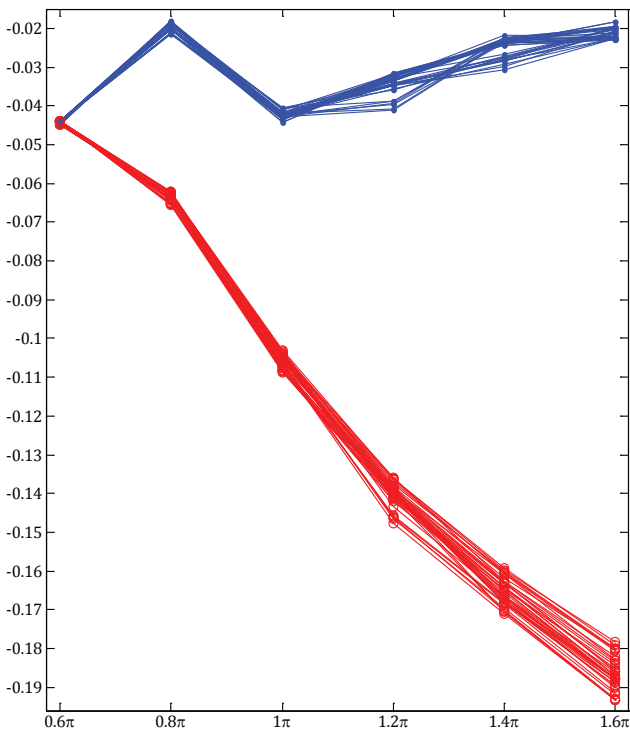


Fig. 5. A bunch of the first player's payoffs at the end of every subinterval (dots) and their cumulative sum (circles) by $M = \overline{15, 20}$ and $J = \overline{15, 20}$

Fig. 6 presenting mixed optimal strategies by $M = \overline{25, 30}$ and $J = \overline{25, 30}$ can be easily compared to Fig. 4. The matter is that, along with the first player's pure optimal strategies on subinterval $[0.4\pi; 0.6\pi)$, the player's mixed optimal strategies on subintervals

$$\left\{ \left\{ [0.2\pi + 0.2\pi i; 0.4\pi + 0.2\pi i] \right\}_{i=2}^5, [1.4\pi; 1.6\pi] \right\} \quad (108)$$

do really converge to the solution of the staircase game. The comparison of more "condensed" payoffs in Fig. 7 to Fig. 5 allows concluding the same. Moreover, here the players' optimal strategies are ε -payoff- $\{M, J\}$ -consistent for

$$\varepsilon = 0.159 \cdot |v_i^*(M, J)| \quad \text{at } i = \overline{1, 6}$$

by every

$$M = \overline{25, 30} \quad \text{and} \quad J = \overline{25, 30}$$

additionally supporting the said.

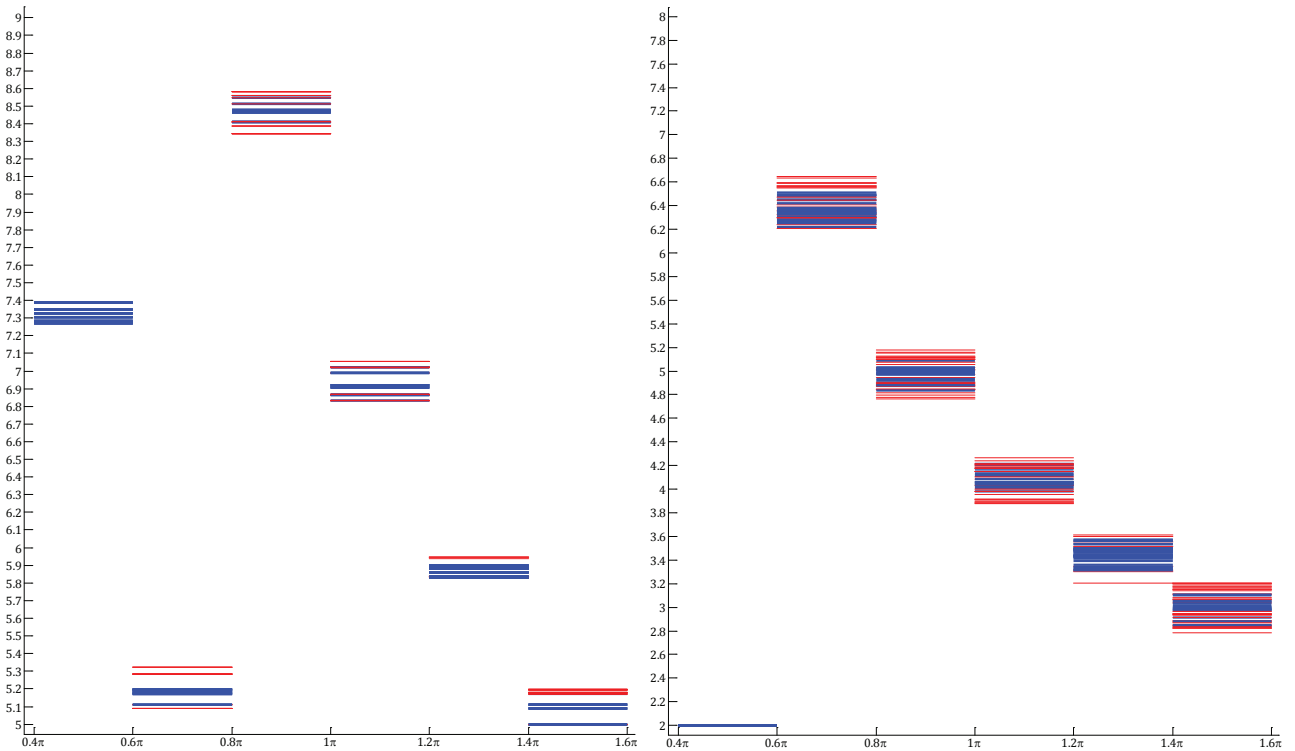


Fig. 6. A bunch of subinterval-wise optimal strategies of the first (left) and second (right) players by $M = \overline{25, 30}$ and $J = \overline{25, 30}$

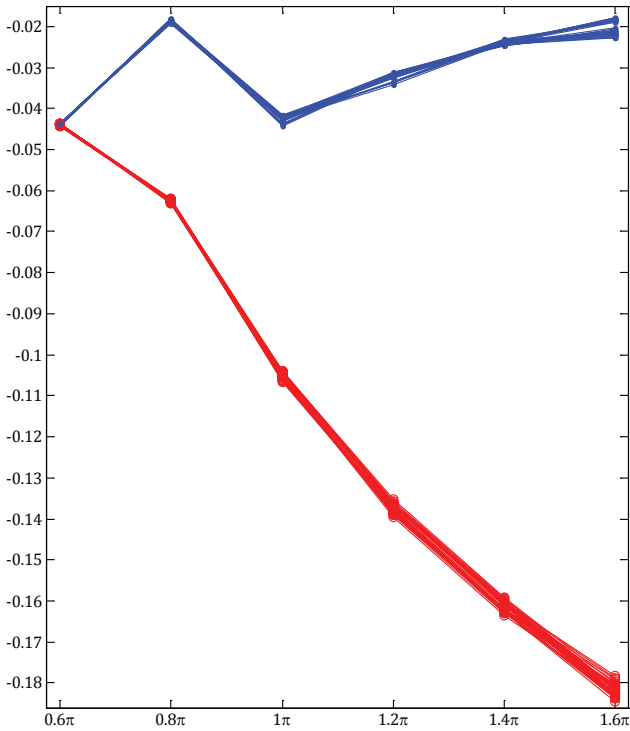


Fig. 7. A bunch of the first player's payoffs at the end of every subinterval (dots) and their cumulative sum (circles) by $M = \overline{25, 30}$ and $J = \overline{25, 30}$

Further thickening the samplings does not change the result much. Along with the first player's pure optimal strategies on subinterval $[0.4\pi; 0.6\pi)$, the scattering of mixed optimal strategies on subintervals (108) by $M = \overline{31, 40}$ and $J = \overline{31, 40}$ (Fig. 8) is slightly less than that in Fig. 6. The "condensation" of payoffs in Figures 7 and 9 are nearly the same.

Although the solution convergence is apparent, the players' optimal strategies are ε -payoff- $\{M, J\}$ -consistent for

$$\varepsilon = 0.202 \cdot |v_i^*(M, J)| \quad \text{at } i = \overline{1, 6}$$

by every

$$M = \overline{32, 39} \quad \text{and} \quad J = \overline{32, 39}.$$

This is an evidence of that the solution convergence reaches its saturation, and further thickening the samplings will not improve the solution approximation nor improve the consistency. Therefore, the approximate solution to the zero-sum staircase game by (94)–(96) and (97)–(100) can be accepted by the independent sampling at both players' with the integers between 25 and 30 (of course, not necessarily identical).

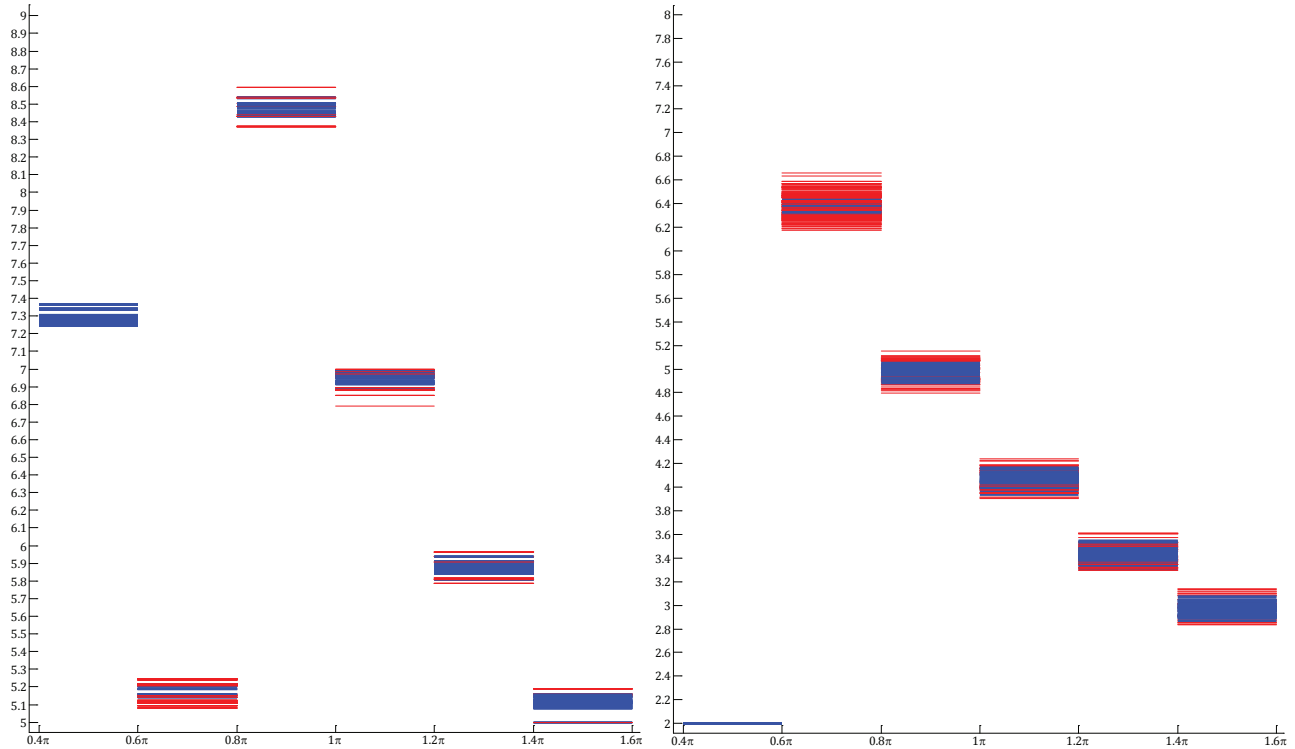


Fig. 8. A bunch of subinterval-wise optimal strategies of the first (left) and second (right) players by $M = \overline{31, 40}$ and $J = \overline{31, 40}$

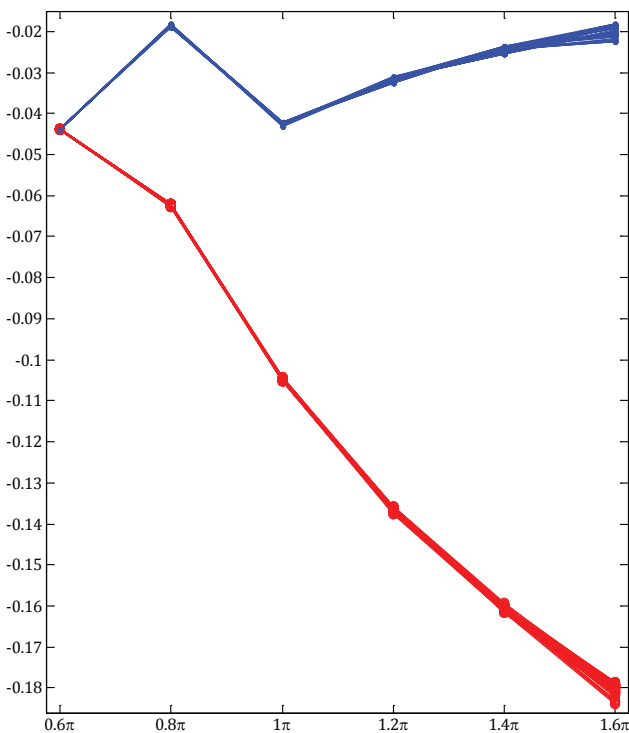


Fig. 9. A bunch of the first player’s payoffs at the end of every subinterval (dots) and their cumulative sum (circles) by $M = \overline{31, 40}$ and $J = \overline{31, 40}$

Discussion of the contribution

Clearly, it would be commonly intractable to straightforwardly solve the sampled staircase game, without considering each subinterval matrix game separately. For instance, by sampling the exemplified game, where each of the players uses 6-subinterval staircase function-strategies, with, say, $M = 25$ and $J = 30$, the resulting $25^6 \times 30^6$ matrix game (in which, e. g., the second player has 729 million pure strategies!) cannot be solved in a reasonable time span. Therefore, solving subinterval matrix games (which are obviously “smaller”) separately and then stacking their solutions is a far more efficient way to obtain an approximate solution of the initial staircase game. The applicability of this method may be limited to the subinterval matrix game size. For instance, the computation time has an exponentially-increasing dependence on the size of the square matrix. Solving matrix games, in which each of the players has at least a few hundred pure strategies, may be time-consuming in applications requiring fast updates of the solution (when the structure of the initial staircase game changes itself).

The (weak) consistency of an approximate solution is a criterion of its acceptability. However, a (weakly) consistent approximate solution may not

exist at appropriately small (tractable) M and J . So, the consistency decomposition into parts by Definitions 2–7 and particularly isolating an ε -payoff consistency by Definition 10 is justified and practically applicable. There are still many open questions, though. First, it is not proved that limits

$$\lim_{M \rightarrow \infty, J \rightarrow \infty} v_i^*(M, J) \quad \forall i = \overline{1, N} \quad (109)$$

exist and they are equal to the respective optimal values of the subinterval continuous games. Second, if limits (109) exist, it is not proved that this is followed by that any approximate solution (44) is ε -payoff- $\{M, J\}$ -consistent for any $M \geq M_*$ and $J \geq J_*$ ($M_* \in \mathbb{N} \setminus \{1\}$, $J_* \in \mathbb{N} \setminus \{1\}$). The inter-influence among the consistency decomposition parts by Definitions 2–7 is also uncertain yet.

The question of a possible reconciliation of the difference of the players’ sampling step selection is indeed that hard. The players can select their samplings simultaneously but not identically. Even if the ranges of function-strategy values are identical and sampling integers M and J are the same (i. e., $M = J$), implying the uniform samplings, a player’s sampling may differ from the other player’s sampling due to eventual inaccuracies in selecting points, as it has been modelled by (102) and (103) with using normal “noise” in the point selection. However, at sufficiently great sampling integers M and J , not necessarily equal, significant changes in M and J are expected not to influence the approximate solution much. Just like in the above-considered example, the player’s optimal strategies converge subinterval-wise and the resulting staircase strategy appears to be an acceptable approximate optimal strategy in the initial staircase game (see Figures 6 and 8). Such a conclusion is made easier by the payoff convergence (see Figures 5, 7 and 9).

Therefore, the presented method is a significant contribution to the antagonistic game theory and its finite approximation supplement. It allows approximately solving zero-sum games with staircase-function strategies in a far simpler manner regardless of the fact that the players may sample their sets of function-strategy values differently [18, 22]. Once the (weak) consistency is confirmed (the respective approximate solution should be at least ε -payoff consistent by Definition 10), the approximate pure-mixed-strategy solution (like those ones of staircase strategies in Figures 6, 8) can be easily implemented and practiced [5, 7, 10, 11, 15, 18, 20].

Conclusion

A zero-sum game played in staircase-function continuous spaces is approximated to a matrix game by sampling the players' pure strategy value sets. Each set is irregularly sampled in its own way so that the resulting samplings may be of different cardinalities and varying densities. While sampled, the requirement of the proper sampling increment (by Definition 1) must be followed – the $S + 1$ points in a 1-incremented sampling must be selected denser than S points.

Owing to Theorem 2, the solution of the matrix game is obtained by stacking the solutions of the “smaller” matrix games, each defined on a subinterval where the pure strategy value is constant. The stack of the “smaller” matrix game solutions is an approximate solution to the initial staircase game. The (weak) consistency of the approximate solution is studied by how much the payoff and optimal situation change as the sampling density minimally increases by the three ways of the sampling increment: only the first player's increment, only the second player's increment, both the players' increment. Thus, the consistency, equivalent to the approximate solution acceptability, is decomposed into the payoff (Definition 2), optimal strategy support cardinality (Definitions 3 and 4), optimal strategy sampling density (Definitions 5 and 6), and support probability consistency (Definition 7).

The most important parts are the payoff consistency and optimal strategy support cardinality (weak) consistency. They are checked in the quickest and easiest way. In addition, it is practically reasonable to consider a relaxed payoff consistency. The relaxed payoff consistency by (91)–(93) means that, as the sampling density minimally increases (in each of the three ways of the sampling increment), the game

optimal value change in an appropriate approximation may grow at most by ε . The weak consistency itself is a relaxation to the consistency, where the minimal decrement of the sampling density is ignored. Therefore, the suggested method of finite approximation of staircase zero-sum games consists in the independent samplings, solving “smaller” matrix games, and stacking their solutions if they are consistent. The finite approximation is regarded appropriate if at least the respective approximate (stacked) solution is ε -payoff consistent (Definition 10).

One can notice that, in staircase game (8) decomposed into games (19), the payoff value depends only on the subinterval length if time t is not explicitly included into the function under the integral in (6). If the subinterval length does not change, every subinterval has the same matrix game. The triviality of the equal-length-subinterval solution is explained by a standstill of the players' strategies. Time variable t explicitly included into (6) means that the players may develop their actions due to the game-modelled system changes (develops) as time goes by.

Finite approximation of games played in staircase-function continuous spaces will be extended and advanced also for the case of non-antagonistic interests of two players sampling their strategy value sets irregularly. An approach to solving the corresponding “smaller” bimatrix games is not straightforwardly deduced from Theorem 2 as the optimality in the matrix game does not have an analogy for the bimatrix game [1, 6, 12, 14, 15]. The independence of the player's sampling step selection may have a deeper incompatibility impact in the bimatrix game case, where multiple and non-equivalent solutions are very often possible, which requires additional reconciliation of the varying profitability.

References

- [1] N. N. Vorob'yov, *Game theory fundamentals. Noncooperative games*. Moscow, USSR: Nauka, 1984.
- [2] N. Nisan, T. Roughgarden, E. Tardos, and V.V. Vazirani, *Algorithmic Game Theory*. Cambridge, MA, USA: Cambridge Univ. Press, 2007, doi: 10.1017/CBO9780511800481.
- [3] M. J. Osborne, *An introduction to game theory*. Oxford, U.K.: Oxford University Press, 2003.
- [4] K. Leyton-Brown and Y. Shoham, “Essentials of game theory: a concise, multidisciplinary introduction,” *Synth. Lect. on Artif. Intell. and Mach. Learn.*, vol. 2, no. 1, pp. 1-88, Jan. 2008, doi: 10.2200/S00108ED1V01Y200802AIM003.
- [5] R. B. Myerson, *Game theory: Analysis of Conflict*. Cambridge, MA, USA: Harvard Univ. Press, 1997.
- [6] N. N. Vorob'yov, *Game theory for economists-cyberneticists*, Moscow, USSR: Nauka, 1985.
- [7] V. V. Romanuke, *Theory of Antagonistic Games*. Lviv, Ukraine: Novyy svit, 2010.
- [8] E. B. Yanovskaya, “Antagonistic games played in function spaces,” *Lithuanian Mathematical Bulletin*, no. 3, pp. 547 — 557, 1967.
- [9] E. B. Yanovskaya, “Minimax theorems for games on the unit square,” *Probability theory and its applications*, no. 9 (3), pp. 554 — 555, 1964.
- [10] T. C. Schelling, *The Strategy of Conflict*. Cambridge, MA, USA: Harvard Univ. Press, 1980.
- [11] H. Moulin, “Extension of two-person zero-sum games,” *J. of Math. Anal. and Appl.*, no. 55 (2), pp. 490 — 507, 1975.
- [12] V. V. Romanuke, “Finite approximation of continuous noncooperative two-person games on a product of linear strategy functional spaces,” *J. of Math. and Appl.*, vol. 43, pp. 123 — 138, 2020, doi: 10.7862/rf.2020.9.

- [13] J. Yang, Y. Chen, Y. Sun, H. Yang, and Y. Liu, "Group formation in the spatial public goods game with continuous strategies," *Phys. A: Stat. Mech. and its Appl.*, vol. 505, pp. 737 — 743, Sep. 2018, doi: 10.1016/j.physa.2018.03.057.
- [14] V. V. Romanuke, "Approximation of unit-hypercubic infinite two-sided noncooperative game via dimension-dependent irregular samplings and reshaping the multidimensional payoff matrices into flat matrices for solving the corresponding bimatrix game," *Comp. Model. and New Technol.*, vol. 19, no. 3A, pp. 7 — 16, 2015.
- [15] V. V. Romanuke and V. G. Kamburg, "Approximation of isomorphic infinite two-person noncooperative games via variously sampling the players' payoff functions and reshaping payoff matrices into bimatrix game," *Appl. Comp. Syst.*, vol. 20, pp. 5 — 14, 2016, doi: 10.1515/acss_2016-0009.
- [16] S. P. Coraluppi, and S. I. Marcus, "Risk-sensitive and minimax control of discrete-time, finite-state Markov decision processes," *Automatica*, vol. 35, no. 2, pp. 301 — 309, Feb. 1999, doi: 10.1016/S0005-1098(98)11253-8.
- [17] S. Rahal, D. Papageorgiou, and Z. Li, "Hybrid strategies using linear and piecewise-linear decision rules for multistage adaptive linear optimization," *Europ. J. of Oper. Res.*, vol. 290, no. 3, pp. 1014 — 1030, 2021, doi: 10.1016/j.ejor.2020.08.054.
- [18] V. V. Romanuke, "Approximation of unit-hypercubic infinite antagonistic game via dimension-dependent irregular samplings and reshaping the payoffs into flat matrix wherewith to solve the matrix game," *J. of Inform. and Org. Sci.*, vol. 38, no. 2, pp. 125 — 143, 2014.
- [19] R. E. Edwards, *Functional Analysis: Theory and Applications*. New York, NY, USA: Holt, Rinehart and Winston, 1965.
- [20] H. Khaloie, A. Abdollahi, M. Shafie-khah, A. Anvari-Moghaddam, S. Nojavan, P. Siano, and J. P. S. Catalao, "Coordinated wind-thermal-energy storage offering strategy in energy and spinning reserve markets using a multi-stage model," *Appl. Energy*, vol. 259, 114168, Feb. 2020, doi: 10/1016/j.apenergy.2019.114168.
- [21] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton, NJ, USA: Princeton University Press, 1944.
- [22] V. V. Romanuke, "Adaptive finite approximation of continuous noncooperative games," *J. of Automat. and Inform. Sci.*, vol. 52, no. 10, pp. 31 — 41, 2020, doi: 10.1615/JAutomatInfScien.v52.i10.20.

В. В. Романюк

СКІНЧЕННА АПРОКСИМАЦІЯ ІГОР З НУЛЬОВОЮ СУМОЮ, ЩО РОЗІГРУЮТЬСЯ У НЕПЕРЕРВНИХ ПРОСТОРАХ СХОДИНКОВИХ ФУНКЦІЙ

Проблематика. Існує відомий спосіб апроксимації неперервних ігор з нульовою сумою, де наближений розв'язок вважається прийнятним, якщо він змінюється мінімально за мінімальної зміни кроку дискретизації. Однак цей метод не можна прямо застосувати до гри з нульовою сумою, що розігрується зі стратегіями у формі сходиноквих функцій. Крім того, слід брати до уваги незалежність вибору гравцем кроку дискретизації.

Мета дослідження. Мета полягає у тому, щоб розробити метод скінченної апроксимації ігор з нульовою сумою, які розігруються у неперервних просторах сходиноквих функцій, беручи до уваги, що гравці, ймовірно, дискретизують множини своїх чистих стратегій самостійно.

Методика реалізації. Для досягнення зазначеної мети формалізується гра з нульовою сумою, в якій стратегії гравців є сходиноквими функціями часу. У такій грі множина чистих стратегій гравця є континуумом сходиноквих функцій часу, і час вважається дискретним. Умови дискретизації множини можливих значень чистої стратегії гравця викладаються так, що гра стає визначеною на добутку скінченних просторів сходиноквих функцій. Загалом, крок дискретизації у кожного гравця різний, і розподіл вибіркових точок (значень функції-стратегії) неоднорідний.

Результати дослідження. Представлено метод скінченної апроксимації ігор з нульовою сумою, які розігруються у неперервних просторах сходиноквих функцій. Метод полягає у нерегулярній дискретизації множини значень чистої стратегії гравця, розв'язуванні матричних ігор меншого розміру, кожна з яких визначена на підінтервалі, де значення чистої стратегії є постійним, й укладанні їхніх розв'язків, якщо вони є узгодженими. Уклад розв'язків матричних ігор меншого розміру є наближеним розв'язком вихідної сходиноквої гри. Досліджується (слабка) узгодженість наближеного розв'язку тим, наскільки змінюється виграш та оптимальна ситуація, коли щільність дискретизації мінімально збільшується трьома способами: лише приріст у першого гравця, лише приріст у другого гравця, приріст в обох гравців. Узгодженість розкладається на узгодженість виграшів, узгодженість потужності спектру оптимальної стратегії, узгодженість щільності дискретизації оптимальної стратегії та узгодженість спектральних імовірностей. З практичної точки зору доцільно розглядати релаксовану узгодженість виграшів.

Висновки. Запропонований метод скінченної апроксимації сходиноквих ігор з нульовою сумою полягає у незалежних дискретизаціях, розв'язуванні матричних ігор меншого розміру за прийнятний проміжок часу та укладанні їхніх розв'язків, якщо вони є узгодженими. Скінченне наближення вважається прийнятним, якщо принаймні відповідний наближений (укладений) розв'язок є узгодженим за ϵ -виграшами.

Ключові слова: теорія ігор; функціонал виграшів; стратегія у формі сходиноквої функції; матрична гра; нерегулярна дискретизація; узгодженість наближеного розв'язку.

Рекомендована Радою
факультету прикладної математики
КПІ ім. Ігоря Сікорського

Надійшла до редакції
15 квітня 2021 року

Прийнята до публікації
14 лютого 2022 року

DOI: 10.20535/kpissn.2021.4.261938

УДК 007: 658.5

П.П. Маслянюк*, І.В. Савчук
КПІ ім. Ігоря Сікорського, Київ, Україна
*corresponding author: mppdom@i.ua

DEVOPS – КОНЦЕПТ І СТРУКТУРНЕ ПРЕДСТАВЛЕННЯ

Проблематика. На сьогодні, поряд з величезною кількістю методів, технологій та засобів підвищення ефективності бізнесу, стрімкого розвитку набуває і застосування концепту DevOps (*Development & Operations*).

Зокрема, застосування в бізнесі концепту DevOps приводить практиків до питання доцільності застосування, формалізації цього процесу, та розробки, як правило, своєї власної системи DevOps зорієнтованої на досягнення визначених значень метрик діяльності Орг.С.

Саме тому автори цієї статті вважають, що тема наукового та прикладного обґрунтування та систематизації DevOps-концепту потребує більш ґрунтовного висвітлення не тільки з точки зору формалізації, але й з точки зору практичного продукування систем DevOps для будь якого виду продуктової діяльності Орг.С.

Мета дослідження. Метою цієї статті є дослідження суті, концепту та дефініції поняття DevOps, визначення сутностей концепту DevOps та формалізація уніфікованої моделі DevOps.

Методика реалізації. Методологія порівняльного аналізу наукових результатів та інженерних практик реалізації DevOps, систематизація класів сутностей концепту DevOps у формі діаграми Вена, графічне моделювання уніфікованої моделі DevOps.

Результати дослідження. Встановлено практичну цінність і результативність впровадження DevOps, відсутність теоретичних досліджень та обґрунтування методів інженерії систем DevOps, окрім прикладів окремих патернів та окремих архітектурних рішень DevOps для ІТ індустрії.

В статті обґрунтовано і запропоновано структурне представлення концепту DevOps у формі діаграми Вена як системи сутностей продукування, підтримки, процесів управління продукуванням і підтримки, та відношень між ними. Ці результати дозволяють обрати теоретичні і прикладні інструменти для формалізації уніфікованої моделі DevOps-систем.

Висновки. Уніфікована модель DevOps-системи у вигляді діаграми компонентів і формалізованих специфікацій інтерфейсів між цими компонентами та уніфікована модель DevOps-системи у вигляді діаграми діяльності цих компонентів дають можливість перейти до розробки обґрунтованих методів інженерії DevOps-систем.

Ключові слова: DevOps, модель DevOps, Системна інженерія DevOps, концепт DevOps, метрики DevOps, процеси DevOps, управління програмним продуктом, колаборація, конкурентна перевага.

Вступ

На сьогодні поряд із величезною кількістю методів, технологій та засобів підвищення ефективності бізнесу стрімкого розвитку набуває і концепт DevOps (*Development & Operations*).

Застосування концепту DevOps і його практична реалізація набувають наразі яскраво вираженого корпоративного характеру, тобто спрямованого на певний вид діяльності та корпоративну культуру тієї чи іншої Організаційної Системи (Орг.С) у самому широкому сенсі розуміння терміна «корпоративна культура».

Зокрема, застосування в бізнесі концепту DevOps приводить практиків до питання доцільності застосування, формалізації цього процесу та розробки, як правило, своєї власної системи DevOps, зорієнтованої на досягнення визначених значень метрик діяльності Орг.С.

У такому контексті власника бізнесу і топ-менеджмент Орг.С цікавлять насамперед повна вартість володіння такою системою DevOps і реальні прибутки від впровадження системи DevOps.

З іншого боку, ринок пропонує не так уже і багато теоретичних та прикладних розробок, на основі яких можна було б реалізувати

Рекомендуємо цитувати цю статтю так: П.П. Маслянюк, І.В. Савчук, «DevOps – Концепт і структурне представлення», *Наукові вісті КПІ*, № 4, с. 39–51, 2021. doi: 10.20535/kpissn.2021.4.261938.

Please cite this article as: P.P. Maslianko, I.V. Savchuk, “DevOps – concept and structural representation”, *KPI Science News*, no. 4, pp. 39–51, 2021. doi: 10.20535/kpissn.2021.4.261938.

фаховий проєкт системи DevOps для конкретної Орг.С.

Ця проблема виникла тому, що наукових публікацій за цим напрямом просто мізерна кількість, а у наявних мова йде переважно про статистичні дослідження впливу DevOps на бізнес у цілому або на окремі метрики Орг.С.

Саме тому автори цієї статті вважають, що тема наукового та прикладного обґрунтування та систематизації DevOps-концепту потребує більш ґрунтовного висвітлення з точки зору не тільки формалізації, але й практичного продукування систем DevOps для будь-якого виду продуктової діяльності Орг.С.

У цій статті досліджуються місце і роль DevOps для будь-якої стадії життєвого циклу продуктів, товарів та послуг, проводиться детальна декомпозиція всіх основних сутностей DevOps, визначаються ролі та вплив кожної з них.

Таким чином, ми спробуємо розкрити всі особливості феномену DevOps не тільки як засобу для організації взаємодії між організаційними одиницями Орг.С, а і як важливого й актуального інструменту для продукування систем DevOps для сучасної ІТ-індустрії.

Постановка задачі

Метою цієї роботи є дослідження суті, концепту та дефініції поняття DevOps, визначення сутностей концепту DevOps та формалізація уніфікованої моделі DevOps.

Суть концепту DevOps

Бачення суті концепту DevOps учасниками ринку та особливостей його застосування показано у багатьох публікаціях. Так, автори статті [1] детально аналізують результати застосування методології DevOps для страхової компанії.

Насамперед автори статті визначають, що *«Сенс DevOps, як висловили респонденти, розглядався як спосіб інтеграції ролей і навичок розробки та операцій ближче один до одного, щоб узгодити стимули ключових ролей, які беруть участь у розробці програмного забезпечення»* [1].

Далі вони вказують, що основна мотивація впровадження методології DevOps в організації — це отримання конкурентної переваги за рахунок застосування й автоматизації бізнес-процесів взаємодії розробників, служб підтримки та менеджменту Орг.С. Автори виділяють дві групи результатів від впровадження DevOps: групу результатів, що безпосередньо стосується конкрет-

них метрик ефективності діяльності Орг.С у цілому, та групу результатів, що стосується метрик оцінювання діяльності самої команди Орг.С.

Окрім результатів, у статті автори [1] вказують і на певні труднощі, пов'язані з впровадженням DevOps.

До першої групи результатів автори відносять такі якісні метрики:

- спільна робота команди, відсутність конфліктів і взаємний контроль, що дає можливість проводити часте/неперервне розгортання нових версій проєкту та впроваджувати нові можливості;

- можливість пришвидшувати процеси розробки та роботу над проєктом у цілому;

- покращення роботи з користувачами результатів розробки та отримання кращого досвіду від користувачів.

До другої групи результатів автори відносять позитивний вплив не тільки на продукти компанії, а й на саму команду:

- учасники команди мають більший рівень автономії та відчувають свою значущість;

- робота в команді мотивує, люди вивчають нові технології;

- учасники команди менше звинувачують один одного, оскільки їхня робота є командною.

Окрім цього, автори визначають і складнощі впровадження DevOps:

- необхідність впроваджувати нові технології та навчати персонал, зокрема інструментам хмарного хостингу, архітектури мікросервісів, а також проведення автоматизованого розгортання та моніторингу;

- необхідність залучати фахівців, знайомих із методологією DevOps.

У статті [2] автори спираються і описують традиційний цикл DevOps і зазначають, що переваги полягають у більшій кількості безперервних розгортань, розширеній автоматизації, кращій відповідності вимогам і загальних покращеннях.

У статті [3] компанія інформує про власний досвід використання методології бізнесом і всередині компанії. Це стаття-звіт щодо методології DevOps від компанії Atlassian яка випускає програмні продукти саме для впровадження методології DevOps (Jira, Bitbucket та інші). Наведені ними статистичні дані чітко вказують на позитивний вплив цієї методології на бізнес за рахунок:

- частіших релізів нових версій проєкту та його швидкого розгортання;

- більшої кількості проєктів, що закінчуються успіхом;

– швидшого налаштування роботи після виявлення несправностей.

У статті вказана і основна парадигма компанії Atlassian стосовно реалізації концепту DevOps – «краще багато маленьких оновлень, але часто, аніж велике та складне оновлення, але досить рідко» [3]. Цю парадигму компанія реалізує за рахунок автоматизації наскрізного бізнес-процесу «ти це зробив, ти це доставив, ти це запустив», тобто DevOps – «це не чиясь окрема робота, це робота усіх, хто працює в компанії» [3].

У звіті Гарвардського університету щодо практик застосування методології DevOps [4] автори зазначають, що наразі одним із ключових показників для бізнесу є якомога швидший запуск продукту з новою функціональністю для кінцевого користувача, але зараз небагато компаній роблять це дійсно добре.

Автори наводять багато інфографіки з приводу того, як оцінюють вплив методології DevOps різні компанії, абсолютна більшість із яких вважають, що вплив є значущим і позитивним.

Автори вказують на те, що впровадження методології – це насамперед проведення змін у самій компанії, управлінні, менеджменті бізнес-процесів і навіть корпоративній культурі.

Корисною є інформація від респондентів про те що, на їх думку, стримує розвиток впровадження DevOps в компанії, зокрема:

- невідповідні організаційні моменти в компанії;
- використання старих (legacy) технологій;
- супротив керівництва до змін;
- обмеженість навичок і правильного способу мислення у робітників;
- необхідність у покращенні безпеки з одночасним підвищенням рівня комунікацій.

Насамкінець автори підкреслюють, що застосування DevOps не є технологічним проектом. На їхню думку, DevOps насамперед полягає у проведенні змін відносно персоналу, організаційної структури, управління продуктивністю і навіть корпоративної культури, тобто всього того, що може фундаментально змінити організацію.

У щорічному огляді від компанії Puppet [5], яка займається випуском програмного забезпечення інструментів DevOps, встановлено, що в компаніях, де впроваджені технології та інструменти DevOps, команди розробки й підтримки мають чітко формалізовані обов'язки та високий рівень автономії щодо виконання цих обов'язків, і головне, вони мають чітко

визначені парадигми взаємодії та канали комунікації між командами [5].

Автори огляду виявили, що види труднощів на шляху впровадження DevOps змінюються від композиції технічних і культурно-організаційних до майже повністю культурних.

Автори огляду вважають, що ключем до максимально успішного рівня впровадження DevOps є застосування єдиної технологічної платформи DevOps для всіх команд компанії. Така платформа регламентує права та обов'язки всіх команд і комунікації між ними.

У статті [6] представлені результати систематичного дослідження, в якому розглянуто основні сфери, теми та цілі NFR (Non-Functional Requirements, нефункціональних вимог) у контексті DevOps, на яких фокусуються розробники при проектуванні DevOps-систем. У статті зазначається, що при дослідженні методології DevOps найчастіше зустрічається такий перелік ключових слів: *Performance, Scalability, Safety, Cross-team integration, Cross-Domain Governance, Regulatory Compliance* та багато інших.

У статті детально проаналізовано та наведено статистику по кожній із частин процесу DevOps. Наведено інформацію про те, що найбільше цікавить технічний, організаційний та бізнесовий склад компаній.

Найважливіші завдання, що можуть бути вирішені за допомогою NFR у DevOps, – це забезпечити безпечну та продуктивну роботу програмного забезпечення (зосередження на технічній якості), сприяти міжкомандній інтеграції експертів із різних галузей (організаційний фокус), з'ясувати відповідність програмного забезпечення нормативним вимогам (фокус на управлінні), використовувати взаємодію клієнта з програмним забезпеченням для ідей щодо перспективних функцій (зосередженість на клієнті), створювати та захищати частку ринку продукту (бізнес-фокус), оптимізувати конвеєр доставки (фокус на розробці) та полегшувати налагодження після розгортання (зосередження на роботі) [6].

Однією з ключових складових DevOps є колаборація. Автори статті [7] дають таке практичне визначення цього терміна: «Колаборація – це здатність людини ефективно робити свій внесок у роботу команди. Це передбачає наполегливість, внесок у знання команди, оцінку внеску інших та здатність до вирішення розбіжностей. Ефективна колаборація передбачає розподіл праці між учасниками, які ведуть активний дискурс, що приводить до об'єднання їхніх зусиль» [7].

У статті [8] автори роблять акцент на тому, що саме в DevOps вважається важливим і найчастіше цитованим: Automated Build, Continuous Deployment, Continuous Integration, Continuous Delivery, Cloud Computing, Source Code Version Control, Automated and Continuous Feedback, Infrastructure as Code.

Автори роблять висновок, що незважаючи на те що в статтях, де згадують DevOps, мова найчастіше йде про програмне забезпечення, все ж не всім вдається вдало застосовувати методологію DevOps, оскільки це не просто технологічний підхід до написання програмного коду, а створення певної моделі культури взаємодії в організації.

Автори наголошують на тому, що результати застосування DevOps проявляються різними формами покращень і перспективами для організацій, які розробляють програмне забезпечення, оскільки його впровадження є чимось трансформуючим для всієї компанії, оскільки залучаються люди, процеси та інструменти. Крім того, впровадження DevOps потребує реструктуризації корпоративної культури, набуття нових навичок, тіснішої співпраці та колаборації команд. Автори підкреслюють, що це нетривіальні задачі, оскільки впровадження DevOps вимагає дотримання низки вимог і дуже добре розробленого та структурованого плану [8].

Важливими є результати, викладені у статті [9], де автори пропонують таксономію рівнів досконалості реалізації DevOps і розглядають декілька патернів організації роботи команд DevOps, починаючи від найновіших, заснованих на досконалій взаємодії між командами Dev і Ops, до більш архаїчних, що не мають достатнього рівня комунікацій між командами.

Таксономія є результатом ґрунтового дослідження, в якому проаналізовано 31 багатонаціональну компанію, що спеціалізуються на розробці програмного забезпечення.

Запропонована таксономія може бути застосована для обґрунтування і реалізації конкретних проєктів DevOps та оцінювання поточного технологічного рівня зрілості DevOps компанії.

Поряд із коротким історичним оглядом методології DevOps, у статті [10] автори роблять огляд важливих складових DevOps і наводять діаграми, що характеризують процес DevOps-розробки. Автори наголошують на тому, що «DevOps – це нова оптимальна практика, якої слід дотримуватися в життєвому циклі розробки програмного забезпечення, щоб збільшити кількість випусків,

надійність, швидкість оновлення, ефективність використання відгуків клієнтів, підвищити якість і безпеку» [10].

Автори підкреслюють, що процес розробки DevOps будь-якої архітектури є нетривіальною задачею, має ітеративний характер, і наголошують, що Хмара є ключовою сутністю DevOps, призначеної для інтеграції команд, безперервної доставки, забезпечення безпеки та збору відгуків від користувачів. Насамкінець автори наголосили, що команда DevOps має бути достатньо кваліфікованою та гнучкою для проведення змін корпоративної культури.

Привертає увагу і стаття [11] із надзвичайно широкою географією і статистикою та висновками про те, як краще запроваджувати DevOps, чи потрібно її взагалі запроваджувати, чи привносить впровадження DevOps якийсь позитивний вплив на результати діяльності компанії.

Автори наводять перелік конкретних практик і принципів Agile, Lean і DevOps, які розглядаються як розширення підходів та інструментів реалізації взаємодії команд DevOps.

Автори встановили, що застосування практик і принципів Agile, Lean і DevOps дає змогу змінювати корпоративну культуру, відстежувати поточний стан стадій життєвого циклу програмного забезпечення та керувати станом розробки засобами автоматизації DevOps.

А застосування SCRUM, Kanban та Continuous Delivery забезпечує доставку, реалізацію процесів управління змінами, управління портфелем послуг (включаючи керування каталогом послуг), керування випуском і розгортанням та керування рівнем обслуговування.

Автори наголошують і на тому, що поряд із високим рівнем підготовки команди DevOps, має бути такий же рівень і керівника DevOps, який повинен працювати на постійній основі [11].

Серед публікацій, що стосуються суті концепту DevOps, є роботи, які описують метрики оцінювання результатів застосування DevOps.

Зокрема, автори статті [6] проаналізували наявні дослідження в галузі DevOps для того, щоб визначити основні вимоги до цієї методології, тобто було проведено аналіз *нефункціональних вимог*. Для цього методологію DevOps розділили на сім доменних областей:

– **Customer Focus Area** (область зосередженості клієнта) – вона концентрується на цінностях, які допомагають покращити досвід клієнта та відслідковувати покращення. Наведемо перелік орієнтованих на клієнта цінностей: *ідеї, які надходять від клієнтів для покращення бізнесу;*

перевірка заявленим вимогам; оцінювання важливості використання клієнтом; частота релізів; пріоритезація циклів розробки, орієнтованих на клієнта.

– **Development Focus Area** (область зосередженості розробки) – в основному тут йдеться про орієнтири на підвищення ефективності процесу розробки. Перелік орієнтованих на розробку цінностей: *вивірність процесу доставки; неперервна інтеграція; неперервна доставка; широкий вибір інструментів розробки; складність задань розробки.*

– **Operation Focus Area** (область зосередженості операцій) – тут йдеться про вимоги, які ставляться до програмного забезпечення і які мають виконуватись під час проведення операцій стосовно нього. Перелік цих вимог: *відслідковування артефактів під час виконання; моніторинг процесу зростання продукту; перевірка продукту після його розгортання.*

– **Governance Focus Area** (область зосередженості управління) – в цій площині розглядаються цінності, які є важливими для процесів управління і які мають бути дотримані процесом DevOps-розробки. Перелік цих цінностей: *забезпечення конфіденційності даних; виконання нормативних вимог; дотримання корпоративної архітектури.*

– **Technical Quality Focus Area** (область зосередженості забезпечення технічної якості) – ця область повністю концентрується на вимогах до продукту, які він має задовольняти, наприклад такі, як час виконання операцій або затрати на використання пам'яті в базі даних. Наведемо розгорнутий перелік цих цінностей і вимог: *зручність використання; стійкість продукту; можливість до залучення нових ресурсів; потужність; використання актуальних фреймворків; безпека даних; можливість розширення продукту; якість поставлених тестів; можливості до подальшого обслуговування продукту; розгортання продукту, орієнтоване на роботу з даними; безпека використання.*

– **Organization Focus Area** (область зосередженості організації) – тут йдеться про цінності та вимоги, які має задовольняти компанія для продукування орієнтованих на бізнес продуктів. Наведемо перелік: *своєчасна реакція на зміну ринку; текучість ресурсів; мотивація для продукування продуктів; комунікація команд в організації; інтеграція команд в організації; можливості для підвищення професійного рівня; охоплення різних доменних областей.*

– **Business Focus Area** (область зосередженості бізнесу) – ця область покриває класичні

вимоги бізнесу, адже його цікавить не те, як буде зроблено продукт, а те, яку цінність він у собі несе. Наведемо перелік таких цінностей і вимог: *середній час до отримання результату; вартість провадження; цінність створеної екосистеми; яку частку ринку займає продукт; актуальність продукту на ринку.*

Таким чином, у цій статті фактично описані класи метрик, необхідних для оцінювання результатів впровадження концепту DevOps.

У статті [12] автор аналізує метрики якості для успішного запуску програмного продукту. Метрики дають можливість визначити, чи буде користь від розробки продукту. Серед таких метрик:

– окупність інвестицій – тобто чи буде приносити продукт дохід від вкладених інвестицій та за який час;

– скільки ідей було згенеровано для того, щоб задовольнити потреби користувача;

– прогнозована комерційна цінність;

– скільки часу потрібно, щоб вийти на ринок;

– яку частину ринку буде займати продукт.

У публікації [13] зазначають важливі метрики оцінювання продукту, який вже працює, наприклад такі як:

– щомісячний регулярний дохід;

– середній дохід за одного користувача;

– вартість залучення нових клієнтів;

– відношення кількості активних клієнтів за день до кількості активних клієнтів за місяць;

– показник утримання клієнтів.

Важливими є і метрики, що оцінюють ефективність роботи команди DevOps. Так, у статті [14] автори вказують на те, що наразі існують чотири найпопулярніші ключові метрики, **Four Key Metrics (FKM)**, для визначення ефективності роботи DevOps і надають фреймворк для автоматизації вимірювання цих метрик. Потрібно зазначити, що ці метрики підходять тільки для IT-продуктів.

1. **Deployment Frequency** (частота розгортань) – ця метрика направлена на те, щоб прискорити зменшення нагромадження в проєкті та прискорювати доставку нових функцій продуктів до кінцевих користувачів.

2. **Lead Time for Change** (час виконання змін) – пояснюється як кількість часу, затраченого на перехід від готового та перевіреного коду до успішно розгорнутого продукту на ринку. Ця метрика вважається важливою тому, що в разі несправності у продукті дає змогу мінімізувати час виправлення проблем.

3. **Time to Restore Service** (час відновлення сервісу) – ця метрика важлива тим, що у великих системах є невідворотним факт того, що будуть з'являтися помилки, і тому дуже важливо мінімізувати час відновлення роботи сервісу.

4. **Change Failure Rate** (рівень змін, які призвели до відмов) – це відсоток нововведень у програмне забезпечення, які призводять до погіршення функціональності або навіть відмови та потребують негайного виправлення (так звані hot-fix).

Таким чином, автори цієї статті пропонують чотири метрики для оцінювання власне роботи команди DevOps з організації розробки та експлуатації програмних продуктів.

На нашу думку, дослідження й аналіз наведених наукових і практичних результатів бачення суті концепту DevOps учасниками ринку дає змогу зробити декілька важливих узагальнювальних висновків.

1. Насамперед звертаємо увагу на множині ключових слів, частина з яких покладена авторами статей в основу онтологічних моделей предметної області DevOps. Зазначимо, що перелік цих ключових слів не є вичерпним!

Ключові слова: DevOps; DevOps system; DevOps model; DevOps adoption; DevOps life cycle; IT product life cycle; DevOps culture; System Design; System Engineering; System Analysis; колаборація; комунікація; неперервна інтеграція; розгортання; автоматизація; моніторинг; тестування; вимірювання метрик; швидкість розробки; гнучкість; конкурентна перевага; cross-team integration; cross-domain governance.

Тут іменники-сутності, іменники-метрики, дієслова-процеси.

2. Команди, тобто організаційні одиниці Орг.С, що задіяні у бізнес-процесах виробництва, підтримки та менеджменту, працюють як складові спільної моделі діяльності Орг.С на всіх стадіях життєвого циклу продуктів, товарів і послуг, у тому числі й розробки програмного забезпечення.

3. Спільна модель діяльності Орг.С накладає і спільну відповідальність за виконання всієї роботи.

4. Спільна модель діяльності Орг.С і спільна відповідальність за виконання всієї роботи надає можливість спільно контролювати та керувати бізнес-процесами діяльності Орг.С.

5. Спільна модель діяльності Орг.С, спільна відповідальність за виконання всієї роботи, спільний контроль і керування бізнес-процесами дає змогу формувати спільне та узгоджене

бачення продукту, його дизайн, конкретну імplementацію та сценарії використання.

6. Спільна модель діяльності Орг.С, спільна відповідальність за виконання всієї роботи, спільний контроль і керування бізнес-процесами, спільне та узгоджене бачення продукту, його дизайн, конкретна імplementація та сценарії використання дають змогу забезпечити можливість для частішого/неперервного розгортання нових версій проекту, проводити аналіз поточних версій і вчасно вносити необхідні зміни.

7. Незважаючи на практичну цінність і результативність впровадження DevOps, привертає увагу відсутність у публікаціях теоретичних досліджень та обґрунтування методів інженерії систем DevOps, окрім прикладів окремих патернів та окремих архітектурних рішень DevOps для IT-індустрії.

8. Більшість авторів підкреслюють, що DevOps є по суті системою автоматизації бізнес-процесів колаборації команд (організаційних одиниць) Орг.С, але при цьому не демонструє DevOps як систему.

9. Оскільки DevOps є по суті системою автоматизації бізнес-процесів колаборації команд (організаційних одиниць) Орг.С, то, на нашу думку, мають бути формалізовані, зокрема, й класи метрик для оцінювання результатів автоматизації бізнес-процесів колаборації.

9.1. Клас метрик оцінювання результатів колаборації бізнес-процесів інженерії продуктів, товарів та послуг [14]:

- частота розгортань (deployment frequency);
- час виконання змін (lead time to change);
- час на відновлення сервісу (time to restore service);
- рівень змін, які призвели до відмов (change failurerate).

9.2. Клас метрик оцінювання результатів колаборації бізнес-процесів підтримки продуктів, товарів і послуг [12, 13]:

- вартість залучення нових клієнтів;
- відношення кількості активних клієнтів за день до кількості активних клієнтів за місяць;
- показник утримання клієнтів;
- кількість згенерованих ідей для задоволення потреб користувача.

9.3. Клас метрик оцінювання результатів колаборації бізнес-процесів управління інженерією та підтримкою продуктів, товарів і послуг на всіх стадіях їх життєвого циклу [12, 13]:

- окупність інвестицій;
- прогнозована комерційна цінність;
- яку частину ринку буде займати продукт;

- скільки потрібно часу, щоб вийти на ринок;
- щомісячний регулярний дохід;
- середній дохід за одного користувача;
- вартість залучення нових клієнтів.

10. На наш погляд, дослідження та розробка методів і моделей інженерії систем DevOps можуть бути корисними і затребуваними для імплементації систем DevOps на всіх стадіях їх життєвого циклу. В будь-якому разі наявність методів і моделей інженерії систем DevOps могла б систематизувати й упорядкувати як процес розробки, так і процеси експлуатації систем DevOps.

DevOps – дефініція та структурне представлення

Для фахової дискусії проведемо аналіз дефініції терміна «DevOps» на основі існуючих поглядів на цю методологію, насамперед у наукових статтях.

Так, автори статті [1] наводять таке визначення терміна «DevOps»:

«DevOps – це набір можливостей інженерних процесів, які підтримуються культурними та технологічними факторами. Можливості визначають процеси, які організація має бути спроможна виконати, в той час як засоби реалізації забезпечують плавність, гнучкість та ефективність роботи».

Це визначення автори підкріплюють переліком елементів множин можливостей і способів їх реалізації інструментами DevOps для Орг.С ІТ-індустрії. Зокрема, до переліку можливостей відносяться:

- колаборативна та неперервна розробка;
- неперервна інтеграція та тестування;
- неперервний випуск і розгортання продукту;
- неперервний моніторинг інфраструктури та оптимізація;
- постійне відслідковування поведінки користувача та реакція на неї;
- відновлення роботи сервісів без затримок;
- неперервне вимірювання показників.

А як способи реалізації цих можливостей зазначають:

- автоматизацію збирання проекту;
- автоматизацію тестування;
- автоматизацію розгортання;
- автоматизацію моніторингу;
- автоматизацію реакції на несправності;
- менеджмент конфігурацій для кодування та інфраструктури;

- автоматизацію проведення вимірювань.

Розглянемо, як автори статті [2] дають визначення терміна «DevOps»:

«DevOps – це підхід, який поєднує традиційні ролі розробників програмного забезпечення та людей, які працюють над покращенням комунікацій у проекті задля підвищення частоти розгортання проекту та підтримки якості програмного забезпечення».

Це визначення є більш конкретним і торкається конкретних понять процесу DevOps-розробки. Серед наборів інструментів, якими оперує DevOps, автори виділяють:

- неперервну інтеграцію;
- неперервне розгортання проекту;
- неперервний моніторинг, насамперед програмного забезпечення;
- контроль версій;
- неперервне тестування.

Автори документу [3] намагаються описувати DevOps максимально просто, щоб донести основну суть цього процесу, і, як вони вважають, у DevOps фактором успіху № 1 є командна робота. В документі наголошується, що DevOps не є роботою якоїсь конкретної людини – це робота всіх працівників, які мають відношення до проекту. Саме за тісної, регулярної взаємодії формується культура праці в колективі, з'являється колективна відповідальність і, як наслідок, бажання брати активну участь в усіх етапах реалізації проекту. Це і декларується у визначенні в їхній інтерпретації:

«**DevOps** – це процес, у якому команди розробки та різних ІТ-операцій працюють разом. Він спрямований на покращення комунікації та колаборації таким чином, щоб процеси розробки, тестування, випуску та використання програмного продукту могли відбуватися швидше та надійніше».

У наступних працях автори описують DevOps з огляду на те, як це поняття характеризують респонденти в опитуваннях і що ці респонденти вважають найбільш вагомими частинами.

У статі [4] автори цікавились, що відзначають люди, які працювали з DevOps. За проведеним опитуванням, 86 % респондентів вважають головною можливістю в DevOps те, що за допомогою цієї методології відбувається пришвидшення розробки та випуску проектів. Компанії вважають, що DevOps надає їм конкурентну перевагу в таких аспектах, як:

- швидкість виходу нового проекту та функцій для кінцевого користувача;
- продуктивність розробки;

- якість розроблюваних продуктів і сервісу в цілому;
- задоволення роботою в командах, що займаються проєктом;
- швидкість впровадження інновацій.

І на основі цього автори доходять висновку, що «DevOps об'єднує команди розробки та операцій для підвищення швидкості та якості продукту, при цьому поєднуючи розробку, інтеграцію, тестування, випуск, моніторинг та управління».

У статті [5] компанія Puppet звітує про DevOps, як «про поняття, яке, можливо, ще не до кінця визначене». Як зазначають автори, ця методологія наразі є дуже популярною серед компаній і зазвичай комбінується з використанням гнучких підходів до розробки, таких як, наприклад, Agile. Окрім цього, автори додають, що залежно від специфіки задачі DevOps може гнучким чином підлаштовуватись під будь-яку предметну область. З урахуванням цього в статті наводяться три визначення того, що таке DevOps, а саме:

- «DevOps – це культурна парадигма та своєрідна об'єднана множина обов'язків, відповідно до цього кожен розробник також несе відповідальність за роботу операцій, а кожен, хто відповідальний за операції, – за участь у розробці»;
- «DevOps – це культурне зміщення, покликане покращити комунікації, колаборацію, інтеграції та автоматизацію в рамках всієї організації»;
- «DevOps – це об'єднання команди, що займається інфраструктурними рішеннями, з командою, що займається розробкою програмного забезпечення, в результаті якого вони разом займаються управлінськими рішеннями та навіть створенням нової інфраструктури».

Зазвичай, коли говорять про DevOps, мають на увазі таку множину сутностей, яку можна представити у вигляді діаграми Венна (див. [15]).

На нашу думку, дуже вдале означення концепту DevOps відносно ІТ-індустрії, наведено в роботі [16]: «DevOps – це набір практик, інструментів та культури, яка автоматизує та інтегрує процеси між командами розробки та командами ІТ-планування. DevOps спеціалізується на тому, що підвищує ефективність, комунікації, співпрацю та автоматизацію роботи ІТ-команд» [1]. Важливо те, що це визначення містить основні класи сутностей, необхідних для розуміння суті концепту DevOps, та суттєво доповнює представлення концепту DevOps у вигляді діаграми Венна (див. [15]).

Зокрема, це сутності інженерії продуктів, товарів і послуг, сутності бізнес-процесів під-

тримки продуктів, товарів і послуг та сутності колаборації бізнес-процесів управління інженерією та підтримкою продуктів, товарів і послуг на всіх стадіях їх життєвого циклу.

Аналізуючи всі ці визначення терміна «DevOps», можна помітити, що більшість авторів відзначають приблизно одне і те ж розуміння сутності концепту DevOps із деякими особливостями, які, швидше за все, обумовлені практичним досвідом авторів цих визначень.

На основі цього огляду виділимо набір слів, якими можна найточніше описати семантику означення концепту DevOps, а саме: *культурна парадигма, колаборація, комунікація, неперервна інтеграція, розгортання, автоматизація, моніторинг, тестування, вимірювання метрик, швидкість розробки, гнучкість, конкурентна перевага*.

Узагальнюючи результати цього короткого системного аналізу сутностей означення концепту DevOps, можна зробити висновок, що мова йде насамперед про модель системи організації процесів розробки та підтримки, принципи якої полягають у застосуванні певних наукових і практичних підходів для об'єднання наявних ресурсів та інтеграції роботи команд розробки й підтримки впровадження.

Концепт DevOps як система інтегрованих міждисциплінарних ресурсів колаборації бізнес-процесів управління інженерією та підтримкою продуктів, товарів та послуг на всіх стадіях їх життєвого циклу

Спираючись на результати системного аналізу сутностей представлення концепту DevOps у вигляді класичної діаграми Венна, дамо авторське визначення концепту DevOps як системи сутностей і відношень між ними. Для цього спочатку формалізуємо складові діаграми Венна з урахуванням класів сутностей предметної області DevOps та дамо їх чітке визначення.

На наш погляд, на підставі результатів аналізу предметної області DevOps, представлення концепту DevOps складається з трьох класів сутностей:

- Множини сутностей інженерії продуктів, товарів та послуг – Entities of Product Development, Set D;
- Множини сутностей підтримки продуктів, товарів та послуг – Entities of Product Support, Set S;
- Множини сутностей управління інженерією та підтримкою продуктів, товарів та послуг – Entities of Support and Engineering Control, Set P.

Далі дамо авторське визначення класів сутностей представлення предметної області концепту DevOps у вигляді діаграми Венна (рис. 1).

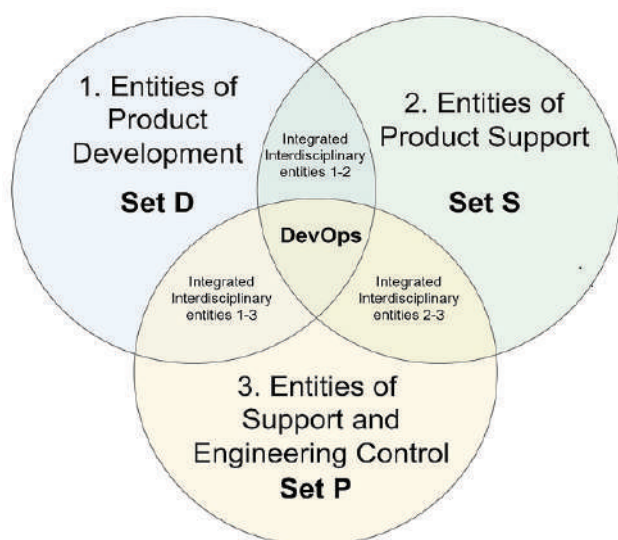


Рис. 1. Представлення класів сутностей концепту DevOps у вигляді діаграми Венна

1. Entities of Product Development, Set D – Множина сутностей інженерії продуктів, товарів та послуг для всіх стадій їх життєвого циклу визначається як множина D , $d \in D$, усіх наявних на ринку матеріальних і нематеріальних сутностей, що можуть бути прямо та/або опосередковано задіяні для інженерії визначеного класу, або об'єкта класу продуктів, товарів та послуг на конкретній стадії їх життєвого циклу.

2. Entities of Product Support, Set S – Множина сутностей підтримки продуктів, товарів та послуг для всіх стадій їх життєвого циклу визначається як множина S , $s \in S$, усіх наявних на ринку матеріальних і нематеріальних сутностей, що можуть бути прямо та/або опосередковано задіяні для підтримки життєдіяльності визначеного класу, або об'єкта класу продуктів, товарів та послуг на конкретній стадії їх життєвого циклу.

3. Entities of Support and Engineering Control, Set P – Множина сутностей управління інженерією та підтримкою продуктів, товарів та послуг для всіх стадій їх життєвого циклу визначається як множина P , $p \in P$, усіх наявних на ринку матеріальних і нематеріальних сутностей, що можуть бути прямо та/або опосередковано задіяні для досягнення мети управління інженерією та підтримкою життєдіяльності продуктів, товарів та послуг визначеного класу, або об'єкта

класу продуктів, товарів та послуг на конкретній стадії їх життєвого циклу.

Наприклад. Development processes – процеси, в ході яких відбуваються розробка, тестування, розгортання проєкту. Включають в себе: кодування; збирання проєкту; неперервне тестування, розгортання, моніторинг, інтеграцію коду; збір артефактів; контроль версій.

Наприклад. Support processes – у цей термін вкладається поняття про те, що проєкт має бути постійно під наглядом, неперервно розгортатися, випускатися, тестуватися, має здійснюватися регулярний контроль за якістю та відповідністю поставленим вимогам. Окрім цього, збирається зворотній зв'язок від користувачів продукту, і формується відповідна реакція на нього.

Наприклад. IT Operations – це найбільш широке поняття в DevOps. Воно включає як постійні корпоративні зустрічі між різними ланками співробітників, так і відповідальність за всі інші види комунікацій на проєкті. Крім цього, операції мають виконуватися із залученням робітників із різними вміннями: наприклад, у розгортанні проєкту мають брати участь як розробники, так і працівники, які займаються тестуванням, та люди, які напряму співпрацюють із бізнесом.

На діаграмі Венна на рис. 1 перетини множин 1-2, 1-3 та 2-3 класів сутностей D , S та P утворюють інтегровані міждисциплінарні ресурси (сутності), виділені за ознакою наявності функціональних відношень між ресурсами (сутностями) різних типів, зокрема Set D, Set S, Set P.

Integrated Interdisciplinary Entities 1-2 (ІІЕ 1-2) – це наявні сутності розробки та підтримки, що дають можливість збільшувати або зменшувати потужності проєкту за потреби.

Integrated Interdisciplinary Entities 2-3 (ІІЕ 2-3) – це наявні сутності, що дають можливість оптимізувати необхідний рівень підтримки продуктів, товарів та послуг на всіх стадіях їх життєвого циклу, постійно випускати нові версії проєкту, неперервно його тестувати та проводити моніторинг.

Integrated Interdisciplinary Entities 1-3 (ІІЕ 1-3) – наявні сутності, що дають можливість оптимізувати процеси продукування продуктів, товарів та послуг на всіх стадіях їх життєвого циклу.

Означення 1. Тут, і надалі у цій статті, інтегровані міждисциплінарні ресурси (сутності) ІІЕ 1-2, ІІЕ 2-3 та ІІЕ 1-3 – це підмножини, що формуються на основі наявності властивостей інтеграції/інтероперабельності між елементами множин D , S та P . А інтеграція/інтероперабельність

елементів множин D , S та P полягає у здатності реалізувати колаборації пар сутностей множин $D i S$, $S i P$, $D i P$.

Тоді для трьох множин – складових DevOps: формалізованих ресурсів (сутностей) інженерії – D ; сутностей інженерії і підтримки життєдіяльності продуктів, товарів та послуг для всіх стадій їх життєвого циклу – S ; сутностей управління інженерією та підтримкою продуктів, товарів та послуг для всіх стадій їх життєвого циклу – P , формалізуємо означення DevOps на основі оновленої діаграми Венна (рис. 1) та наявності властивостей інтеграції/інтероперабельності між елементами цих множин.

Означення 2. Концепт DevOps може бути формалізований як множина трійок елементів міждисциплінарних ресурсів (сутностей) із трьох множин ресурсів (сутностей): сутностей інженерії продуктів, товарів та послуг для всіх стадій їх життєвого циклу – D , сутностей підтримки продуктів, товарів та послуг для всіх стадій їх життєвого циклу – S та сутностей управління інженерією та підтримкою продуктів, товарів та послуг для всіх стадій їх життєвого циклу – P таких, що мають спільну властивість/ознаку, яка забезпечує інтеграцію/інтероперабельність між елементами $d \in D$, $s \in S$, $p \in P$ і встановлює відповідності колаборації C_{dspd} і C_{dpsd} між елементами цих множин.

Тобто

$$\begin{aligned} \text{DevOps} = \{ & d, s, p, d \mid \forall d \in D, \forall s \in S, \\ & \forall p \in P, (d, s, p, d) \in C_{dspd} \} \cup \\ & \cup \{ d, p, s, d \mid \forall d \in D, \forall p \in P, \\ & \forall s \in S, (d, p, s, d) \in C_{dpsd} \}, \end{aligned} \quad (1)$$

або

$$\begin{aligned} \text{DevOps} = & C_{dspd} \cup C_{dpsd} = \\ = & \{ \text{кортеж}(D, S, P, D, G_{dspd}) \cup \\ \cup & \text{кортеж}(D, P, S, D, G_{dpsd}) \}, \end{aligned} \quad (2)$$

де: G_{dspd} і G_{dpsd} – графіки/діаграми/матриці відповідностей колаборації C_{dspd} і C_{dpsd} .

Таким способом ми обґрунтовуємо структурне представлення DevOps оновленої діаграми Венна за допомогою наявності властивості/ознаки відповідності, що забезпечує інтеграцію/інтероперабельність між елементами $d \in D$, $s \in S$, $p \in P$ і встановлює відповідності колаборації C_{dspd} і C_{dpsd} між елементами множин ресурсів D , S і P оновленої діаграми Венна (рис. 1).

Така декомпозиція класів сутностей D , S і P та узагальнене системне представлення кола-

борації елементів множин цих сутностей представлення концепту DevOps (1, 2) дає змогу систематизувати існуючі означення та формалізувати дефініцію концепту «DevOps» як колаборацію сутностей інженерії, сутностей підтримки та сутностей управління інженерією та підтримкою продуктів, товарів та послуг на всіх стадіях їх життєвого циклу.

Така декомпозиція і систематизація дає нам глибше уявлення того, що таке концепт DevOps, які завдання він вирішує і, в той же час, показує, наскільки широкою є область знань, що застосовується для організації виробництва продуктів, товарів та послуг протягом їх життєвого циклу.

Це дає нам можливість узагальнити поняття «концепт DevOps» та сформулювати означення «DevOps-система».

Означення 3. Концепт DevOps – міждисциплінарна методологія управління колаборацією сутностей інженерії та підтримки життєдіяльності продуктів, товарів та послуг на всіх стадіях їх життєвого циклу.

Виходячи з усього викладеного вище, можна формалізувати поняття DevOps як системи та перейти до поширення цього поняття не тільки на розробку як процес, а і на активний бізнес-процес на кожній зі стадій життєвого циклу продуктів, товарів та послуг.

Означення 4. DevOps-система, або система DevOps, – це множина сутностей і відношень між ними, необхідна і достатня для забезпечення колаборації сутностей інженерії та підтримки життєдіяльності продуктів, товарів та послуг на всіх стадіях їх життєвого циклу за заданих умов і обмежень.

Тоді на підставі результатів такого системного аналізу можна формалізувати структурне та динамічне представлення уніфікованої моделі системи DevOps.

Зокрема, структурне представлення уніфікованої моделі DevOps-системи у вигляді діаграми компонентів у нотації UML показано на рис. 2.

Опишемо кожен з інтерфейсів уніфікованої моделі DevOps-системи у вигляді діаграми компонентів:

Dec-to-Dev – інтерфейс, який слугує для того, щоб передати інформацію про цілі бізнесу, бажаний продукт, вимоги до нього та поставлені задачі розробникам.

Dec-to-Ops – інтерфейс передавання інформації про цілі бізнесу, бізнес-вимоги, бажаний отриманий результат, важливі сфери бізнесу та поставлені задачі команді операцій.

DevOps Система

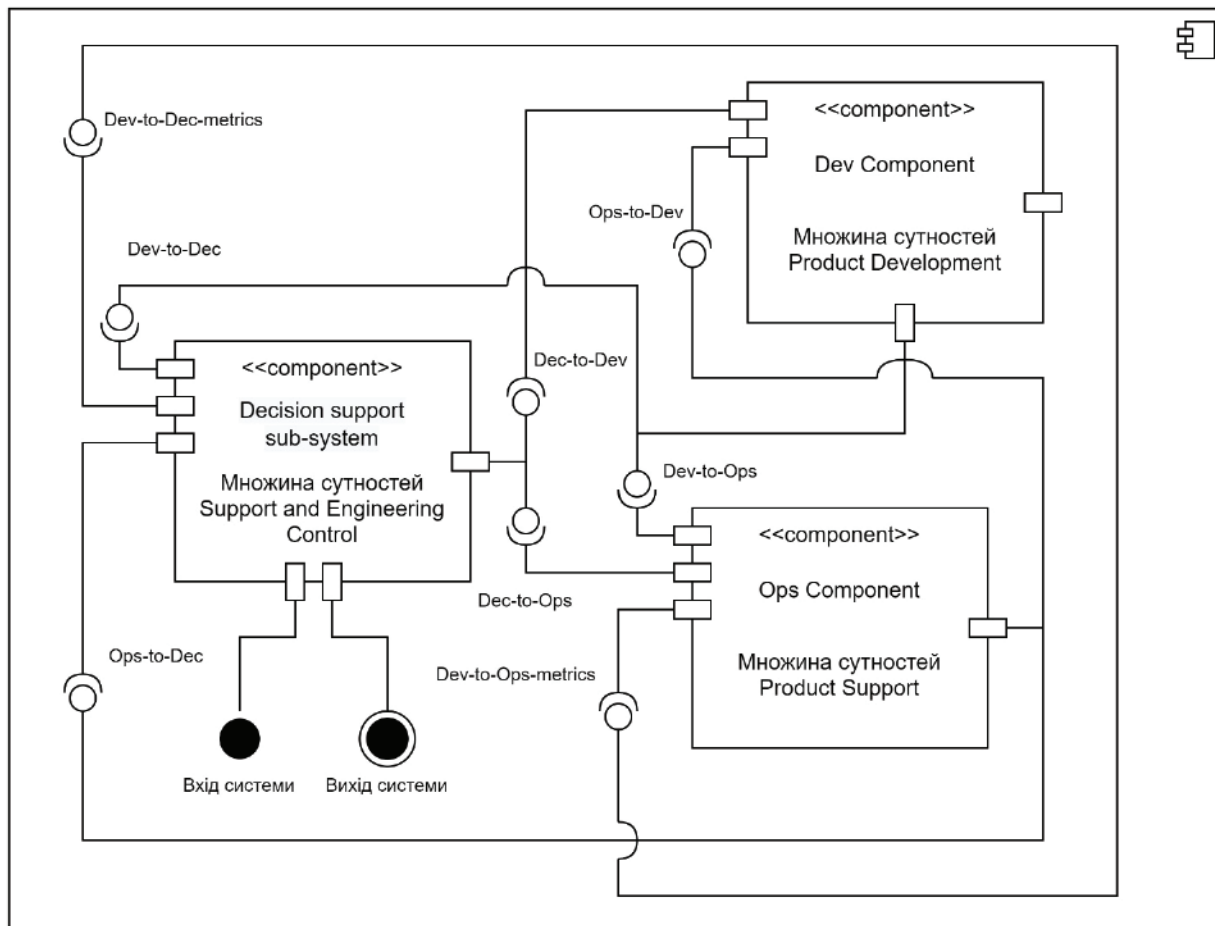


Рис. 2. Уніфікована модель DevOps-системи. Діаграма компонентів у нотатції UML

Dev-to-Dec – інтерфейс передавання інформації від команди розробників до системи управління, а саме: прогрес, отримані результати, затребувані ресурси тощо.

Dev-to-Ops – інтерфейс передавання інформації від команди розробників до команди операцій, а саме: отримані артефакти, програмний код для валідації, документація, зворотній зв'язок тощо.

Ops-to-Dev – інтерфейс передавання зворотного зв'язку та нових задач від команди операцій до команди розробників.

Ops-to-Dec – інтерфейс передавання інформації про прогрес на проекті, затребувані ресурси, уточнення бізнес-вимог системі управління.

Dev-to-Dec-metrics – інтерфейс передавання інформації про впровадження продуктів, їх потужність, зовнішній вигляд продуктів, відповідність продукту реаліям ринку.

Ops-to-Dec-metrics – інтерфейс передавання інформації про впровадження продуктів, відсоток покриття бізнес-вимог, відсоток покриття вимог щодо функціональності продукту, основні показники ефективності (Key Performance Indicators, KPI) продукту.

А динамічне представлення уніфікованої моделі DevOps-системи у вигляді діаграми діяльності в нотатції UML має вигляд, показаний на рис. 3. Діаграма діяльності формалізує колаборацію компонентів уніфікованої моделі DevOps-системи.

Уніфікована модель DevOps-системи у вигляді діаграми компонентів і формалізованих специфікацій інтерфейсів між цими компонентами та уніфікована модель DevOps-системи у вигляді діаграми діяльності цих компонентів дають можливість перейти до розробки обґрунтованих методів інженерії DevOps-систем.

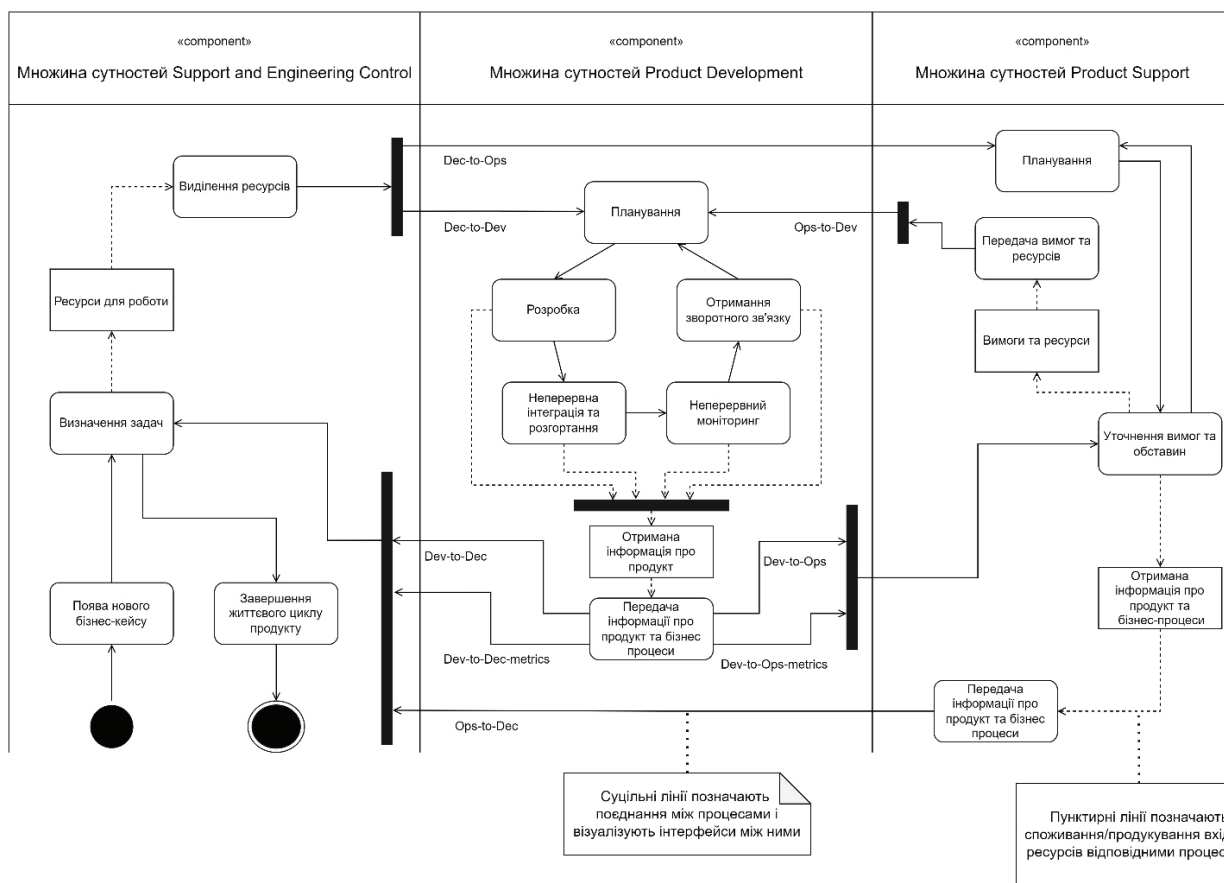


Рис. 3. Уніфікована модель DevOps-системи. Діаграма діяльності у нотатції UML

Висновки

1. У результаті аналізу наукової літератури та кращих інженерних практик встановлено, що незважаючи на практичну цінність і результативність впровадження DevOps, у наявних публікаціях відсутні теоретичні дослідження та обґрунтування методів інженерії систем DevOps, окрім прикладів окремих патернів та окремих архітектурних рішень DevOps для ІТ-індустрії.

Більшість авторів підкреслюють, що DevOps є по суті системою автоматизації бізнес-процесів колаборації команд (організаційних одиниць) Орг.С, але при цьому не розглядають DevOps як систему.

2. У статті обґрунтовано і запропоновано структурне представлення концепту DevOps у формі діаграми Венна як системи сутностей продукування, підтримки, процесів управління продукуванням і підтримки та відношень між ними. Ці результати дають можливість вибрати теоретичні та прикладні інструменти для формалізації уніфікованої моделі DevOps-систем.

3. Запропонована уніфікована модель DevOps-системи може бути застосована для розробки обґрунтованих методів інженерії DevOps-систем, призначених для управління колаборацією сутностей інженерії та підтримки життєдіяльності продуктів, товарів та послуг на всіх стадіях їх життєвого циклу.

References

[1] M. Senapathi, J. Buchan, H. Osman, “DevOps capabilities, practices, and challenges: Insights from a case study”, in *Proc. 22nd Int. Conf. Eval. Assess. Software Eng.*, 2018.

[2] P. Jha, R. Khan, “A review paper on DevOps: Beginning and more to know”, *Int. J. Comp. Applic.*, vol. 180, no. 48, pp. 16–20, 2018. doi: 10.5120/ijca2018917253

- [3] DevOps 101 with Atlassian [Online]. Available: <https://www.atlassian.com/blog/wp-content/uploads/devops-101-atlassian.pdf>
- [4] “Competitive Advantage through DevOps”, Harvard Business Review Analytic Service, Res. Rep., 2019.
- [5] “2021 Puppet State of DevOps”, CircleCI, Report, 2021.
- [6] P. Haindl, R. Plusch, “Focus areas, themes, and objectives of non-functional requirements in DevOps: A systematic mapping study”, in *46th Euromicro Conf. Software Eng. Advanc. Applic.*, Portoroz, Slovenia, pp. 394–403, 2020. doi: 10.1109/SEAA51224.2020.00071
- [7] C. Scoular, D. Duckworth, J. Heard, D. Ramalingam, “Collaboration: Definition and structure”, The Australian Council for Educational Research, report, 2020.
- [8] M.L. Pedra, M.F. Silva, L.G. Azevedo, “DevOps adoption: Eight emergent perspectives”, ArXiv abs/2109.09601, 2021.
- [9] D. Lypez-fernández, J. Díaz, J.G. Martín, J.E. Pérez, B. González-Prieto, “DevOps team structures: Characterization and implications”, ArXiv abs/2101.02361, 2021.
- [10] M. Gokarna, R. Singh, “DevOps: A historical review and future works”, in *Proc. 2021 Int. Conf. Comp. Commun. Intell. Syst.*, Greater Noida, India, pp. 366–371, 2021. doi: 10.1109/ICCCISS1004.2021.9397235
- [11] K. Maroukian, S.R. Gulliver, “Leading DevOps practice and principle adoption”, ArXiv abs/2008.10515, 2020.
- [12] N. Bhuiyan, “A framework for successful new product development”, *J. Indust. Eng. Manag.*, vol. 4, no. 4, 2011. doi: 10.3926/jiem.334.
- [13] R.G. Oliveira, “Lean and performance measuring – Developing a new performance measurement framework to fit lean”, M.S. thesis, Universidade de Porto, 2018.
- [14] M. Sallin, M. Kropp, C. Anslow, J.W. Quilty, A. Meier, “Measuring software delivery performance using the four key metrics of DevOps”, in *Agile Processes in Software Engineering and Extreme Programming. XP 2021. Lecture Notes in Business Information Processing*, vol 419, P. Gregory, C. Lassenius, X. Wang, P. Kruchten, eds. Cham: Springer, 2021. doi: 10.1007/978-3-030-78098-2_7
- [15] R. Pietrantuono, A. Bertolino, G.D. Angelis, B. Miranda, S. Russo, “Towards continuous software reliability testing in DevOps”, in *2019 IEEE/ACM 14th Int. Workshop on Automation of Software Test*, Montreal, QC, Canada, 2019. doi: 10.1109/AST.2019.00009
- [16] “What is DevOps?” [Online]. Available: <https://www.atlassian.com/devops>
- [17] A. Kosyakov, S.M. Beamer, S.J. Seymour, W.N. Sweet, *Systems Engineering. Principles and Practice*. Moscow, Russia: DMK Press, 2014, 624 p.
- [18] D.K. Hitchins, *Systems Engineering. A 21st Century Systems Methodology*. Wiley, 2007, 528 p.

P.P. Maslianko, I.V. Savchuk

DEVOPS – CONCEPT AND STRUCTURAL REPRESENTATION

Problems. Today, along with a huge number of methods, technologies and means of improving business efficiency, the concept of DevOps (Development & Operations) is also gaining rapid development.

In particular, the application of the DevOps concept in business leads practitioners to the question of the feasibility of application, formalization of this process, and usually the development of their own DevOps system focused on achieving the specified values of the metrics of the organizational structure.

Research goal and objectives. The purpose of this article is to study the essence, concept and definition of the DevOps concept, define the entities of the DevOps concept and formalize the unified DevOps model.

Methods of research. Methodology of comparative analysis of scientific results and engineering practices of DevOps implementation, systematization of entity classes of the DevOps concept in the form of a Venn diagram, graphic modeling of a unified DevOps model.

Research results. The practical value and effectiveness of DevOps implementation, the lack of theoretical research and justification of DevOps systems engineering methods, apart from examples of individual patterns and individual architectural solutions of DevOps for the IT industry, have been established.

The article substantiates and proposes a structural representation of the DevOps concept in the form of a Venn diagram as a system of entities of production, support, management of production and support processes, and relationships between them. These results make it possible to choose theoretical and applied tools for the formalization of a unified model of DevOps systems.

Conclusions. The unified model of the DevOps system in the form of a diagram of components and formalized specifications of the interfaces between these components and the unified model of the DevOps system in the form of a diagram of the activities of these components make it possible to proceed to the development of well-founded engineering methods of DevOps systems.

Keywords: DevOps, DevOps model, DevOps system engineering, DevOps concept, DevOps metrics, DevOps processes, software product management, collaboration, competitive advantage.

Рекомендована Радою
факультету прикладної математики
КПІ ім. Ігоря Сікорського

Надійшла до редакції
22 жовтня 2021 року

Прийнята до публікації
14 лютого 2022 року

DOI: 10.20535/kpissn.2021.4.261849

УДК 608 (075.8)

Ю.М. Кузнецов, Г. Сінмін, О.В. Самойленко*

КПІ ім. Ігоря Сікорського, Київ, Україна

*corresponding author: o.samoilenko@kpi.ua

ПЕРЕДУМОВИ ВИКОРИСТАННЯ СИСТЕМНО-МОРФОЛОГІЧНОГО ПІДХОДУ ТА ТЕОРІЇ ФРАКТАЛІВ ПРИ СТВОРЕННІ ЛЕЩАТ ДЛЯ ОБ'ЄКТІВ СКЛАДНОЇ ФОРМИ

Проблематика. На металорізальних верстатах, у механоскладальному виробництві й побуті широко використовують лещата різних конструкцій, які переважно мають постійну структуру й один кінематичний ланцюг від двигуна (чи ручного приводу) до затискних елементів – губок. Їх переважна більшість пристосована для затиску циліндричних предметів (деталей) і предметів із плоскопаралельними сторонами. Водночас у механоскладальному виробництві, а також під час ремонтних робіт часто виникає необхідність затиску предметів сферичних, клиноподібних, конічних, неправильної форми.

Мета дослідження. Запропонувати морфологічну множину універсальних лещат з альтернативами нових ознак стану, виду виконання і рухомості затискних елементів, а також скласти морфологічні формули лещат.

Методика реалізації. Розроблюється морфологічна модель лещат, яка враховує досягнення теорії фракталів і дає змогу спрогнозувати нові технічні рішення універсальних лещат на рівні винаходів.

Результати дослідження. Запропоновані морфологічні формули універсальних лещат з морфологічної множини, яка містить альтернативи нових ознак стану, виду виконання і рухомості затискних елементів, а також передумови використання теорії фракталів для моделювання процесу затиску деталей складної форми.

Висновки. Використовуючи запропонований підхід, можна спрогнозувати нові технічні рішення універсальних лещат на рівні винаходів для виготовлення, дослідження та подальшої комерціалізації.

Ключові слова: лещата, морфологічний підхід, теорія фракталів, моделювання процесу затиску.

Вступ

На металорізальних верстатах, у механоскладальному виробництві й побуті широко використовують лещата різних конструкцій, які переважно мають постійну структуру й один кінематичний ланцюг від двигуна (чи ручного приводу) до затискних елементів – губок [1]–[3]. Наявність одного кінематичного ланцюга знижує ККД і потребує додаткових витрат енергії (за механізованого приводу) чи мускульної роботи працівника (за ручного приводу). Функціональні можливості лещат обмежені тим, що їх переважна більшість пристосована для затиску циліндричних предметів (деталей) і предметів із плоскопаралельними сторонами. Крім того, у механоскладальному виробництві, а також під час ремонтних робіт постає проблема затиску

предметів сферичних, клиноподібних, конічних, неправильної форми.

Актуальним стає розроблення нових універсальних лещат із перемінною структурою, яка має два й більше кінематичних ланцюгів та дає можливість суттєво розширити їхні функціональні можливості (адаптація до об'єкта будь-якої форми та широкодіапазонність) [4]. Це сприятиме зменшенню кількості оснащення та технологічних операцій, скороченню часу на підготовку виробництва та полегшенню праці людей під час виробничих процесів і ремонтних робіт. Над створенням таких лещат працюють фахівці, однак окреслена науково-технічна задача нині повністю не розв'язана та потребує подальших досліджень, що визначає їх актуальність.

Аналіз попередніх досліджень [4]–[7] дає можливість зробити такі висновки:

Рекомендуємо цитувати цю статтю так: Ю.М. Кузнецов, Г. Сінмін, О.В. Самойленко, “Передумови використання системно-морфологічного підходу та теорії фракталів при створенні лещат для об'єктів складної форми”, *Наукові вісті КПІ*, № 4, с. 52–57, 2021. doi: 10.20535/kpissn.2021.4.261849.

Please cite this article as: Yu. N. Kuznietsov, X. Gao, and O. V. Samoilenko “Prerequisites for using the system-morphological approach and fractal theory in the creation of vises for complex shape objects”, *KPI Science News*, no. 4, pp. 52–57, 2021. doi: 10.20535/kpissn.2021.4.261849.

1. Уже використовують лещата з двома кінематичними ланцюгами, зокрема з диференціальними гвинтовими передачами. В цих конструкціях перший кінематичний ланцюг призначений для вибірки зазору між заготовкою та губками за допомогою ручного чи механізованого приводу. Другий кінематичний ланцюг, де використовують гідравлічні й електромеханічні перетворювачі, призначений для остаточного затиску деталей. Але досі немає теоретичних основ створення слюсарних і машинних лещат такого типу.

2. В КПІ ім. Ігоря Сікорського, з використання системно-морфологічного підходу та сучасних методів пошуку нових технічних рішень виконувались дослідження, метою яких є створення лещат із заданими чи покращеними функціональними можливостями [3]–[7]. Ці дослідження присвячені силовим і жорсткісним характеристикам машинних лещат, а також використанню плинних і сипких середовищ у затискних елементах. Водночас ще немає єдиного методологічного підходу до проектування універсальних лещат.

3. Відомо небагато конструкцій лещат універсального призначення [8], в яких без зміни затискних губок можливо затискати заготовки різної форми, хоча у США давно запатентовано лещата [9] із ручним приводом і гвинтовою передачею за принципом, що використовується в математичній теорії фракталів.

4. Розроблення та проектування універсального оснащення вимагають нових методологічних підходів [10], [11], прогресивних методів структурно-схемного синтезу за пошуку нових рішень і створення й використання ефективних математичних моделей й алгоритмів за допомо-

гою комп'ютерів. Напрямок поглиблених досліджень може бути пов'язаний не тільки з методологією науково-технічної творчості, а й з теорією фракталів – геометричних фігур, що складаються з частин, кожна з яких, незалежно від розміру, виглядає як вся фігура [8].

Термін “фрактал” запровадив у 1975 р. французький математик Бенуа Мандельброт [13], [14]. Польський математик Вацлав Серпінський ще раніше запропонував різні геометричні структури, [15] по суті фрактальні, які можна успішно використовувати в техніці, наприклад, при створенні антен [16].

Розглянемо деякі структури Серпінського, що підходять для моделювання та створення універсальних лещат типу “фрактальні лещата”.

На рис. 1 показано так званий трикутник Серпінського, отриманий унаслідок чотирьох циклів побудови. Беруть рівносторонній трикутник, який є основою для побудови фрактала (ітерація 0); цей трикутник розбивають на чотири рівносторонніх трикутники, а в отриманій фігурі видаляють центральну частину (ітерація 1); вказану операцію повторюють для кожного з отриманих менших трикутників (ітерації 2–4).

На рис. 2 показано квадрат Серпінського, який будується таким чином. Береться квадрат (ітерація 0), кожна сторона квадрата ділиться на три рівні частини, а весь квадрат відповідно на дев'ять однакових квадратиків. Центральний квадрат вирізається (ітерація 1). Кожний з восьми квадратів, що залишилися теж діляться на дев'ять квадратів, центральний з яких вирізається (ітерація 2). Процедура повторюється для тих квадратів, що залишилися і т. д.



Рис. 1. Трикутник Серпінського

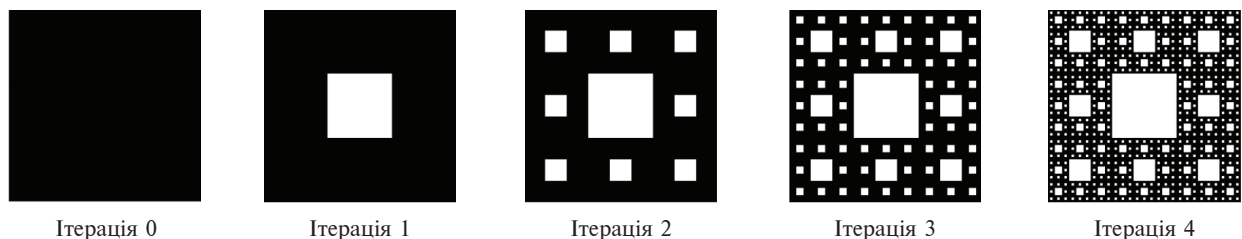


Рис. 2. Квадрат Серпінського

Постановка задачі

Метою роботи є: запропонувати морфологічну множину універсальних лещат з урахуванням досягнень у теорії фракталів й альтернативами нових ознак стану, виду виконання та рухомості затискних елементів; спрогнозувати технічні рішення універсальних лещат на рівні винаходів для виготовлення, дослідження та подальшої комерціалізації.

Результати досліджень

Для пошуку та прогнозування нових рішень лещат розроблено морфологічну модель – таблицю, що може розширюватися з появою нових ідей (табл. 1).

Для збереження конфіденційності цю морфологічну модель можна подати як морфологічну множину – повну морфологічну матрицю [13, 14] $M_{лщ} = M_{пз} \times M_{пт} \times M_{зе} \times M_{оз}$, де $M_{пз}$ – матриця приводу затиску (ПЗ) (ознака 1); $M_{пт}$ – матриця перетворювача (ПТ) (ознака 2); $M_{зе}$ – матриця затискних елементів (ЗЕ) з ознаками стану (3), виконання (4), виду рухомості (5); $M_{оз}$ – матриця об'єкту затиску з ознаками кількості (6) та форми (7):

$$M_{лщ} = \begin{matrix} | & 2.1 \\ | & 2.2 \\ 1.1 | & 2.3 \\ | & 2.4 \\ | & 2.5 \\ 1.2 | & 2.6 \\ | & 2.7 \\ | & 2.8 \end{matrix} \times \begin{matrix} | & 3.1 & 4.1 & 5.1 \\ | & 3.2 & 4.2 & 5.2 \\ | & 3.3 & 4.3 & 5.3 \\ | & 3.4 & 4.4 & 5.4 \\ | & 3.5 & 4.5 & 5.5 \\ | & 4.6 & & \end{matrix} \times \begin{matrix} | & 6.1 & 7.1 \\ | & 6.2 & 7.2 \\ | & 6.3 & 7.3 \end{matrix}$$

Це дає загальну кількість варіантів рішення:

$$N_{лщ} = 6 \times 8 \times 5 \times 6 \times 5 \times 3 \times 3 = 64800.$$

Для зменшення поля пошуку та зручного показу моделі запропоновано усічену матрицю з суттєвим зменшенням варіантів можливих рішень:

$$M_{лщ} = \begin{matrix} | & 2.1 \\ | & 2.2 \\ 1.1 | & 2.3 \\ | & 2.4 \\ | & 2.5 \end{matrix} \times \begin{matrix} | & 4.1 & 5.1 \\ | & 4.2 & 5.2 \\ | & 4.3 & 5.3 \\ | & 4.4 & 5.3 \\ | & 4.5 & \\ | & 4.6 & \end{matrix} \times \begin{matrix} | & 6.1 & 7.1 \\ | & 7.2 \end{matrix}$$

Таблиця 1. Морфологічна модель лещат

1. Привід затиску (ПЗ)	2. Перетворювач (ПТ)	Затискні елементи (ЗЕ)			Об'єкт затиску (ОЗ)	
		3. Стан	4. Виконання	5. Вид рухомості	6. Кількість	7. Форма
1.1. Ручний	2.1. Гвинтовий	3.1. Твердотілі цільні плоскі	4.1. Одномісне симетричне	5.1. Поступовий	6.1. Один	7.1. Правильна
1.2. Електро-механічний	2.2. Важільний	3.2. Твердотілі призматичні	4.2. Одномісне несиметричне	5.2. Поступово-обертовий	6.2. Кілька в один ряд	7.2. Складна
1.3. Гідрравлічний	2.3. Спіральний	3.3. Твердотілі складені	4.3. Багатомісне симетричне в один ряд	5.3. Поступово-обертовий-обертовий	6.3. Кілька в декілька рядів	7.3. Комбінована
1.4. Пневматичний	2.4. Плунжерний	3.4. Плинні	4.4. Багатомісне симетричне в кілька рядів	5.4. Поступово-обертовий-обертовий багатократно	–	–
1.5. Електромагнітний	2.5. Клиновий	3.5. Рідинні	4.5. Багатомісне несиметричне в один ряд	5.5. Обертовий	–	–
1.6. Гідро-(пневно-) механічний	2.6. Зубчастий	3.6. Сипкі	4.6. Багатомісне несиметричне в кілька рядів	–	–	–
–	2.7. Пружний	–	–	–	–	–
–	2.8. Комбінований	–	–	–	–	–

Тоді загальна кількість варіантів рішення буде такою:

$$N_{лш}^1 = 2 \times 5 \times 3 \times 6 \times 3 \times 1 \times 2 = 1080.$$

Для ручного ПЗ, який застосовується переважно в слюсарних лещатах (альтернатива 1.1), й електромеханічного ПЗ, який використовується в машинних лещатах (альтернатива 1.2), з розповсюдженим гвинтовим ПТ (альтернатива 2.1) запишемо кортежі морфологічної множини як морфологічні формули (сполучення альтернатив кожної ознаки) для ОЗ простої форми (альтернатива 7.1) $X_1 - X_8$ і складної форми типу “шарнірний палець кермової тяги автомобіля” (альтернатива 7.2) X_9 і X_{10} (рис. 3):

$$X_1 = (1.1) - (2.1) - (3.1 - 4.1 - 5.1) - (6.1) - (7.1) - \text{поз. “а”}$$

$$X_2 = (1.1) - (2.1) - (3.1 - 4.2 - 5.1) - (6.1) - (7.1) - \text{поз. “б”}$$

$$X_3 = (1.1) - (2.1) - (3.2 - 4.1 - 5.1) - (6.1) - (7.1) - \text{поз. “в”}$$

$$X_4 = (1.1) - (2.1) - (3.2 - 4.2 - 5.1) - (6.1) - (7.1) - \text{поз. “г”}$$

$$X_5 = (1.1) - (2.1) - (3.3 - 4.1 - 5.1) - (6.1) - (7.1) - \text{поз. “д”}$$

$$X_6 = (1.1) - (2.1) - (3.3 - 4.2 - 5.1) - (6.1) - (7.1) - \text{поз. “е”}$$

$$X_7 = (1.1) - (2.1) - (3.3 - 4.1 - 5.3) - (6.1) - (7.1) - \text{поз. “ж”}$$

$$X_8 = (1.1) - (2.1) - (3.3 - 4.2 - 5.3) - (6.1) - (7.1) - \text{поз. “з”}$$

$$X_9 = (1.1) - (2.1) - (3.3 - 4.1 - 5.2) - (6.1) - (7.2) - \text{поз. “і”}$$

$$X_{10} = (1.1) - (2.1) - (3.3 - 4.1 - 5.3) - (6.1) - (7.2) - \text{поз. “к”}$$

Жирним шрифтом позначено альтернативи, які відрізняються від варіанта X_1 .

Ефективність розробленої морфологічної моделі (див. табл. 1) підтверджено тим, що в просторі множини рішень знаходиться й патент США № 1059545 (рис. 4) [12], який описується морфологічною формулою: 1.1–2.1–3.3–4.1–5.4–6.1–7.2, а поз. 1, 2, 3, 4 відповідають поворотним ЗЕ з осями повороту на кути $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, у поз. 5 розташований гвинт 6 для поступального переміщення ЗЕ.

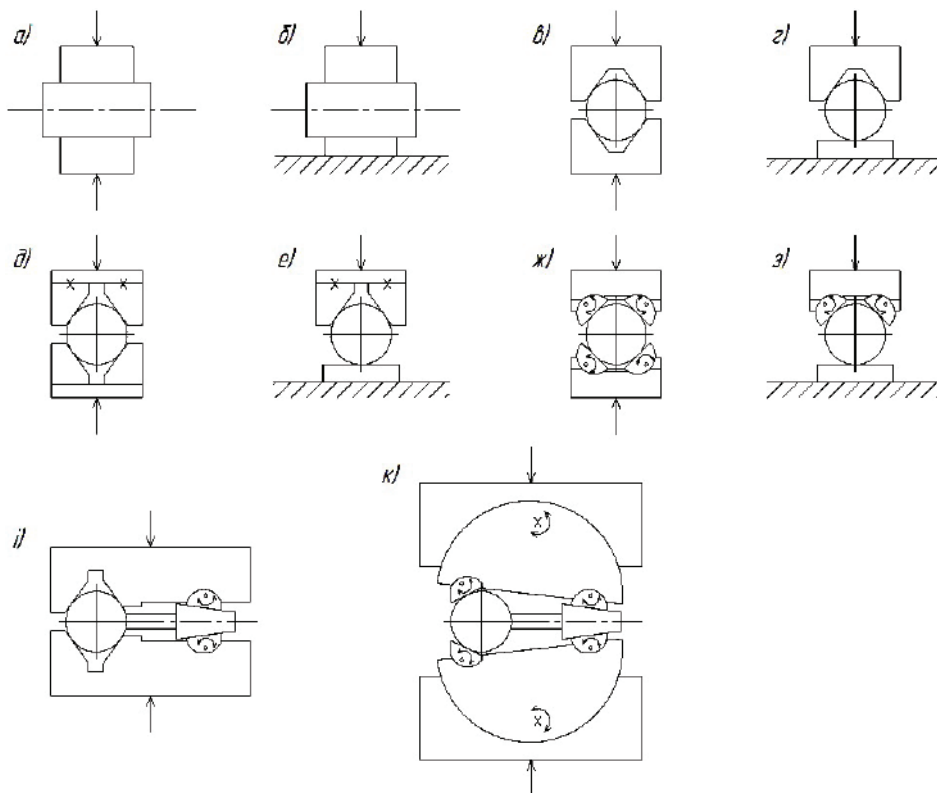


Рис. 3. Варіанти загику об'єктів простої (а–з) і складної (і, к) форми

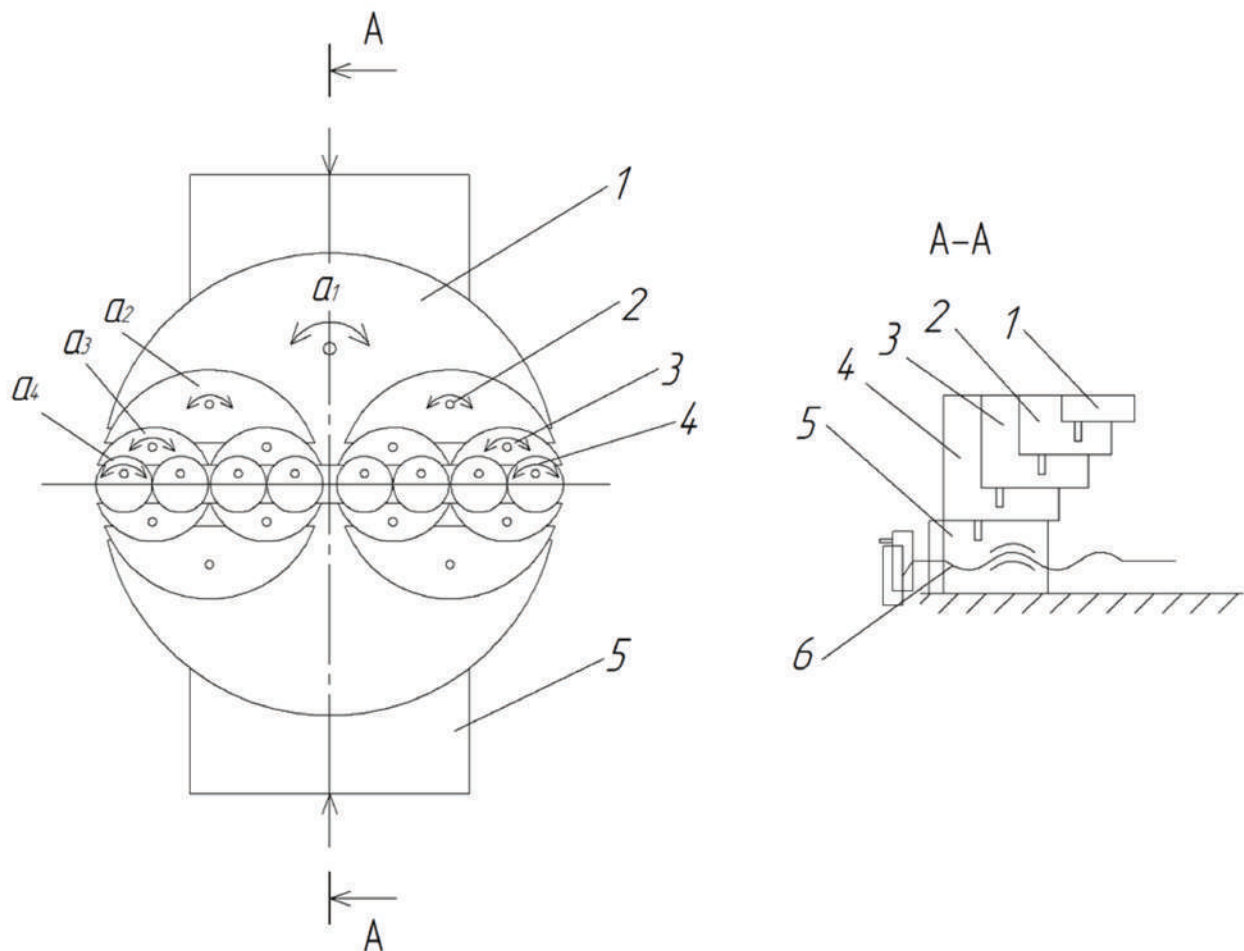


Рис. 4. Схема фрактальних лещат за патентом США № 1059545 із ручним приводом затиску і гвинтовим перетворювачем

Висновки

Уперше запропоновано морфологічну множини конструкцій універсальних лещат з урахуванням досягнень у теорії фракталів й альтернативами нових ознак стану, виду виконання

та рухомості затискних елементів, а також із написанням морфологічних формул конструкцій лещат. Використовуючи описаний підхід, можна спрогнозувати технічні рішення стосовно універсальних лещат на рівні винаходів для виготовлення, дослідження та подальшої комерціалізації.

References

- [1] *Machine vices with manual and power drives. General technical conditions*, GOST 16518-87, 1988.
- [2] A. I. Astakhov et al., *Machine tools. A handbook in two volumes*, vol. 1. B. N. Vardashkin and A. A. Shatilov, Eds., Moscow, USSR:, Mashinostroenie, 1984.
- [3] Yu. M. Kuznetsov, B. A. Kryzhanivskiy, and T. O. Khamuiela, "System-morphological approach in the creation of clamping devices for clamping prismatic workpieces", *Naukovi pratsi KNTU*, no. 5, 2004.
- [4] Yu. N. Kuznetsov, T. A. Khamuiela, and P. M. Nedelcheva, "Creation of clamping devices with multiple kinematic chains for prismatic workpieces", in *Works Int. Sci. Conf. UNITECH-05*, Gabrovo, Bulgaria, 2005, pp. 127–132.
- [5] T. O. Khamuiela, "Experimental studies of elastic-force characteristics of machine vices for metal-cutting machines", *Visnyk NTUU "KPI", ser. "Mashynobuduvannia"*, pp. 115–121, 2006.
- [6] Yu. M. Kuznetsov, T. O. Khamuiela, and S. P. Yakovenko, "Vice," UA Patent 15583, July 17, 2006.
- [7] Yu. M. Kuznetsov, T. O. Khamuiela, and M. M. Panchenko, "Vice," UA Patent 21917, April 10, 2007.

- [8] M. W. Thring and E. R. Laithwaite, *How to Invent*, Moscow, USSR: Mir, 1980.
- [9] P. K. Kunze, "Device for obtaining intimate contact with, engaging, or clamping bodies of any shape", U.S. Patent 1059545, April 22, 1913.
- [10] V. M. Odrin, *Method of morphological analysis of technical systems*. Moscow, USSR: VNIPI, 1989.
- [11] Yu. M. Kuznetsov and G. Xingmin, "Using a system-morphological approach when searching for new vices to clamp objects of correct and complex shape", in *Pratsi XXIII Mizhnarodnoi naukovo-tekhnichnoi konferentsii ASPHP*, Kyiv, Ukraine: HAV, 2021, pp. 140–145.
- [12] M. V. Pratsiovytyi, *Fractal approach in studies of singular distributions*, Kyiv, Ukraine: Vydavnytstvo NPU imeni M. P. Drahomanova, 1998.
- [13] B. Mandelbrot, *Fractals: Form, chance, and dimension*, San-Francisco: W. H. Freeman and Co., 1977.
- [14] B. Mandelbrot, *The Fractal Geometry of Nature*, San-Francisco: W. H. Freeman and Co., 1982
- [15] W. Sierpiński, "Sur une courbe dont tout point est un point de ramification," *Compt. Rendus Acad. Sci. Paris*, vol. 160, pp. 302–305, 1915.
- [16] V. M. Vishnevsky, A. I. Lyakhov, Portnoy S. L., and, I. V. Shakhnovich, *Broadband wireless information transmission networks*, Moscow, Russia: Tekhnosfera, 2005.

Yu. M. Kuznetsov, G. Xingmin, O. V. Samoilenko

PREREQUISITES FOR USING THE SYSTEM-MORPHOLOGICAL APPROACH AND FRACTAL THEORY IN THE CREATION OF VICES FOR COMPLEX SHAPE OBJECTS

Background. On metal-cutting machines, in mechanical assembly production and everyday life, vices of various designs are widely used, which, as a rule, have a constant structure and one kinematic chain from the engine (or manual drive) to the clamping elements – jaws. The vast majority of them are adapted for clamping cylindrical objects (parts) and objects with plane-parallel sides. At the same time, in mechanical assembly production, as well as during repair work, it often becomes necessary to clamp spherical, wedge-shaped, conical, and irregularly shaped objects.

Objective. The purpose of the paper is to propose a morphological set of universal vices taking into account the achievements in the theory of fractals and alternatives to new features of the state, type of execution and mobility of clamping elements, and to come up with morphological formulae of vices.

Methods. A morphological model of the vices is being developed, which takes into account the achievements of the fractal theory and makes it possible to predict new technical solutions for universal vices at the level of inventions.

Results. The morphological formulae of universal vices from the morphological set which contains alternatives of new signs of a condition, a type of execution and mobility of clamping elements, and also prerequisites for using the fractal theory for clamping process modelling of complex shape details are offered.

Conclusions. Using the proposed approach, it is possible to predict new technical solutions of universal vices at the level of inventions for the manufacture, research and subsequent commercialization.

Keywords: vices, morphological approach, fractal theory, clamping process modelling.

Рекомендована Радою
Навчально-наукового інституту матеріалознавства та зварювання
КПІ ім. Ігоря Сікорського

Надійшла до редакції
8 жовтня 2021 року

Прийнята до публікації
14 лютого 2022 року

DOI: 10.20535/kpissn.2021.4.263195
УДК 621.762+621.793

Д.О. Ремізов*, А.М. Степанчук, Т.О. Соловійова, П.І. Лобода, С.В. Завадюк

КПІ ім. Ігоря Сікорського, Київ, Україна

*corresponding author: remizz@ua.fm

ВИКОРИСТАННЯ САМОФЛЮСІВНИХ СПЛАВІВ ЯК ДЖЕРЕЛА ЛЕГУВАЛЬНИХ ЕЛЕМЕНТІВ ПІД ЧАС ОТРИМАННЯ ПОРОШКОВИХ ВИРОБІВ КОНСТРУКЦІЙНОГО ПРИЗНАЧЕННЯ З НИЗЬКОЛЕГОВАНОГО ЗАЛІЗА

Проблематика. Однією з проблем виготовлення порошкових конструкційних матеріалів на основі заліза, легованих малою кількістю легувальних елементів (0,5–5,0 мас.%), є отримання гомогенного сплаву в кінцевому матеріалі. Рівномірність розподілу легувальних елементів суттєво визначає кінцеві властивості матеріалу. Використання суміші порошків складників для отримання таких сплавів зазвичай не розв'язує проблеми отримання гомогенних сплавів. Тож перспективним є використання порошків сплавів, які вміщують легувальні елементи, як джерела легувальних елементів.

Мета дослідження. Вивчити закономірності формування структури та властивостей матеріалів на основі сплавів заліза, отриманих із використанням самофлюсівних сплавів (СФС) як джерела легувальних елементів.

Методика реалізації. Застосовано методики змішування–розмелювання у планетарних млинах, вивчення процесів змочування розплавами СФС заліза та взаємодії між ними методом лежачої краплі, растрової мікроскопії для вивчення структури матеріалів і розподілу елементів у них, вимірювання механічних властивостей (твердості).

Результати дослідження. Встановлено засади визначення оптимальних умов отримання матеріалів: час розмелювання–змішування (4 хв), тиск пресування вихідних заготовок (600 МПа), температура спікання (1200 °С), час спікання (понад 80 хв). Їх застосування сприяє отриманню виробів з практично стовідсотковою щільністю матеріалу та високим ступенем гомогенності.

Висновки. Використання СФС як джерела легувальних елементів, а також методів готування вихідної суміші змішуванням–розмелюванням у планетарному млині сприяє отриманню матеріалів із високим ступенем гомогенності внаслідок інтенсифікації процесів гомогенізації та спікання та, як наслідок, і підвищенню їх механічних характеристик.

Ключові слова: залізо; легувальні елементи; розмелювання; пресування; спікання; структура матеріалу; твердість.

Вступ

Однією з проблем виготовлення конструкційних матеріалів на основі заліза, легованого малою кількістю легувальних елементів (0,5–5,0 мас.%), є отримання гомогенного сплаву у кінцевому матеріалі. Рівномірність розподілу легувальних елементів істотно визначає кінцеві властивості матеріалу, особливо характеристики міцності. Але за використання як вихідних матеріалів порошків складників матеріалу в кількостях, менших за 5 мас. % (особливо меншій за 1 мас. %),

утворення гомогенних сплавів практично неможливе в разі отримання вихідної шихти для пресування заготовок виробу механічним змішуванням. У цьому випадку важко забезпечити рівномірний контакт між складниками сплаву, що вимагає значного збільшення часу спікання для забезпечення гомогенізації сплаву.

Цю проблему можна частково розв'язати, застосовувавши для виготовлення виробів порошки готових сплавів. Ці порошки можна отримати розпилюванням розплавів, сумісним відновленням оксидів металів, дифузійним наси-

Рекомендуємо цитувати цю статтю так: Д.О. Ремізов, А.М. Степанчук, Т.О. Соловійова, П.І. Лобода, С.В. Завадюк, “Використання самофлюсівних сплавів, як джерела легувальних елементів при отриманні порошкових виробів конструкційного призначення з низьколегованого заліза”, *Наукові вісті КПІ*, № 4, с. 58–68, 2021. doi: 10.20535/kpissn.2021.4.263195.

Please cite this article as: D.O. Remizov, A.M. Stepanchuk, T.O. Soloviova, P.I. Loboda, and S.V. Zavadiuk, “The use of self-fluxing alloys as a source of alloying elements in producing powder products of construction purpose from low-alloyed iron”, *KPI Science News*, no. 4, pp. 58–68, 2021. doi: 10.20535/kpissn.2021.4.263195.

ченням, насиченням із точкових джерел [1]–[5]. Однак використання вказаних методів для отримання порошків сплавів, особливо малолегованих, не завжди гарантує бажаний результат і є економічно не вигідними.

Деякою мірою проблему створення більшої площі контакту між складниками сплаву за малої кількості легувальної складової частини можна розв'язати, застосовувавши хімічне змішування [6], [7]. У цьому випадку легувальну добавку під час змішуванні вихідних порошків вводять у вигляді розчину її солі. Під час змішування з одночасним випаровуванням розчинника сіль рівномірно осаджується на поверхні частинок основного матеріалу сплаву (у нашому випадку заліза) та відновлюється (розкладається). Тим самим значно збільшується площа контакту між складовими сплаву й утворюються активні джерела насичення основного матеріалу легувальними елементами, як це відбувається й за отримання сплавів насиченням із точкових джерел [1], [3]. Однак використання цієї технології теж не завжди ефективно через дефіцитність розчинних солей певних елементів, їх невідповідність вимогам відновлюваності чи розкладання та високу вартість, що значно знижує економічну ефективність методу.

Одним із варіантів отримання конструкційних матеріалів зі сплавів на основі заліза, що містять легувальні елементи в кількості, меншій за 1–5 мас. %, є використання порошків сплавів, які містять необхідні легувальні елементи. Так, перспективними є порошки СФС [1], [8]. Їх легко отримати розпилюванням розплавів. Вони можуть бути на основі нікелю, заліза та кобальту [1], [2] та містити вуглець, силіцій, бор, хром, молібден, мідь тощо. За необхідності до їх складу також можна вводити інші елементи. Особливістю СФС є їх відносно низька температура плавлення (залежно від складу – 920–1180 °С). Тому за спікання конструкційних порошкових матеріалів на основі заліза за температури 1200–1250 °С вони плавимуться й утворюватимуть рідку фазу, яка за змочування розплавом СФС заліза розтікатиметься поверхнею частинок і проникатиме між ними, що значно збільшуватиме контактну поверхню між ними й так інтенсифікуватиме процеси сплавоутворення. У цьому разі слід чекати утворення (за інших рівних умов) більш гомогенних сплавів і підвищення характеристик їх міцності.

Постановка задачі

Метою роботи є отримати матеріали на основі заліза, легованого СФС на основі нікелю марки АН9 [2] як джерела легувальних елементів; дослідити вплив умов змішування вихідних порошків на структуру та деякі властивості матеріалів.

Результати та їх обговорення

Однією з бажаних вимог до сплавів, які використовують як джерело легувальних елементів, є утворення рідкої фази під час спікання виробів за їх участю. У цьому разі значно прискорюються дифузійні процеси, що лежать в основі гомогенізації (сплавоутворення) кінцевого матеріалу. Але ці процеси відбуватимуться тільки у випадку змочування легувальної добавки твердої фази (у нашому випадку карбонільного заліза) розплавом. Тому в роботі досліджено змочуваність розплавом СФС заліза. Змочування вивчалось методом лежачої краплі [9], [10]. Досліджувався вплив на крайовий кут змочування температури (1100 °С; 1150 °С, 1200 °С і 1250 °С) і часу змочування (1, 2, 3, 4, 5 с).

Результати досліджень наведено на рис. 1.

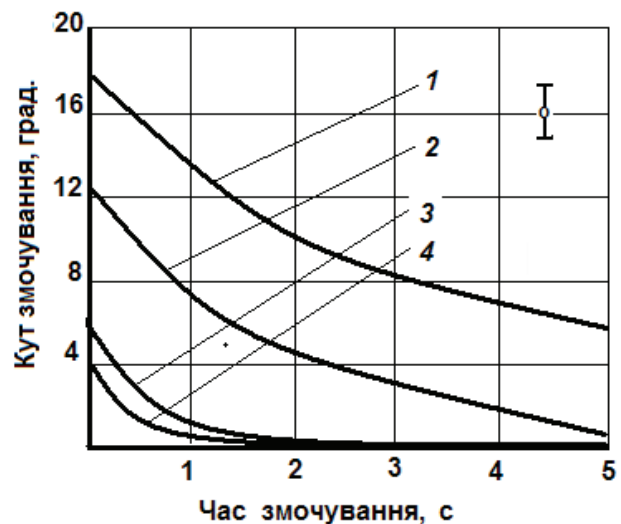


Рис. 1. Залежність кута змочування розплавом СФС поверхні заліза від часу змочування за різних температур (°С): 1 – 1100 °С; 2 – 1150 °С; 3 – 1200 °С; 4 – 1250 °С

Як видно з рис. 1, розплав СФС практично у всіх досліджених умовах змочує залізо (кут змочування менший за 90°). У нашому випадку він значно менший за 90° і змінюється залежно від

умов від 18 до 0°. Кут змочування зменшується зі збільшенням температури та часу змочування. За температур 1200 °С і більше відбувається повне змочування вже за 1–2 с, тобто практично миттєво. Отримані результати узгоджуються з даними роботи [11], у якій досліджено процеси змочування розплавами СФС заліза. Показано, що наявність у залізі інших легувальних елементів, особливо карбідоутворюючих, дещо збільшує кут змочування, але за певних умов, які застосовують у практиці порошкової металургії при отриманні порошкових конструкційних матеріалів на основі заліза, він прагне до нуля.

Тож із погляду відповідності вимогам змочування використання СФС як джерела легувальних елементів є перспективним.

Іншим питанням, яке необхідно розв'язувати під час отримання конструкційних матеріалів на основі заліза, легованого малою кількістю легувальних елементів, є визначення умов спікання спресованих заготовок. У нашому випадку ставиться задача не тільки отримати вироби з максимальною щільністю, від якої суттєво залежать їхні механічні властивості [12], а й забезпечити рівномірний розподіл легувальних елементів у кінцевому виробі, який також має значний

вплив на їх експлуатаційні властивості. Для визначення цих умов необхідно знати температуру та час спікання заготовок, спресованих із суміші вихідних порошків. Аналіз джерел [6], [7], [13] дав змогу визначити температуру та час спікання порошкових матеріалів на основі заліза для отримання виробів із заданими властивостями (щільністю). Але у нашому випадку необхідно поряд із максимальною щільністю отримати й рівномірний розподіл легувальних елементів у матеріалі з використанням їх як джерела СФС. Тому важливо з'ясувати поряд із часом спікання час гомогенізації матеріалу. Останній можна оцінити за результатами визначення кінетики взаємодії заліза з розплавами СФС. Задля цього процеси взаємодії розплаву СФС із компактным залізом досліджено з використанням зразків із чистого (армко) заліза, на поверхню яких розміщували зразок із СФС. Нагрівали зразки за температур, вищих за температуру плавлення останнього, 1100 та 1200 °С. Для дослідження впливу температури та часу на ступінь взаємодії між розплавом СФС і залізом вивчено мікроструктуру зони взаємодії та розподіл і вміст елементів у її межах. Результати наведено на рис. 2–4.

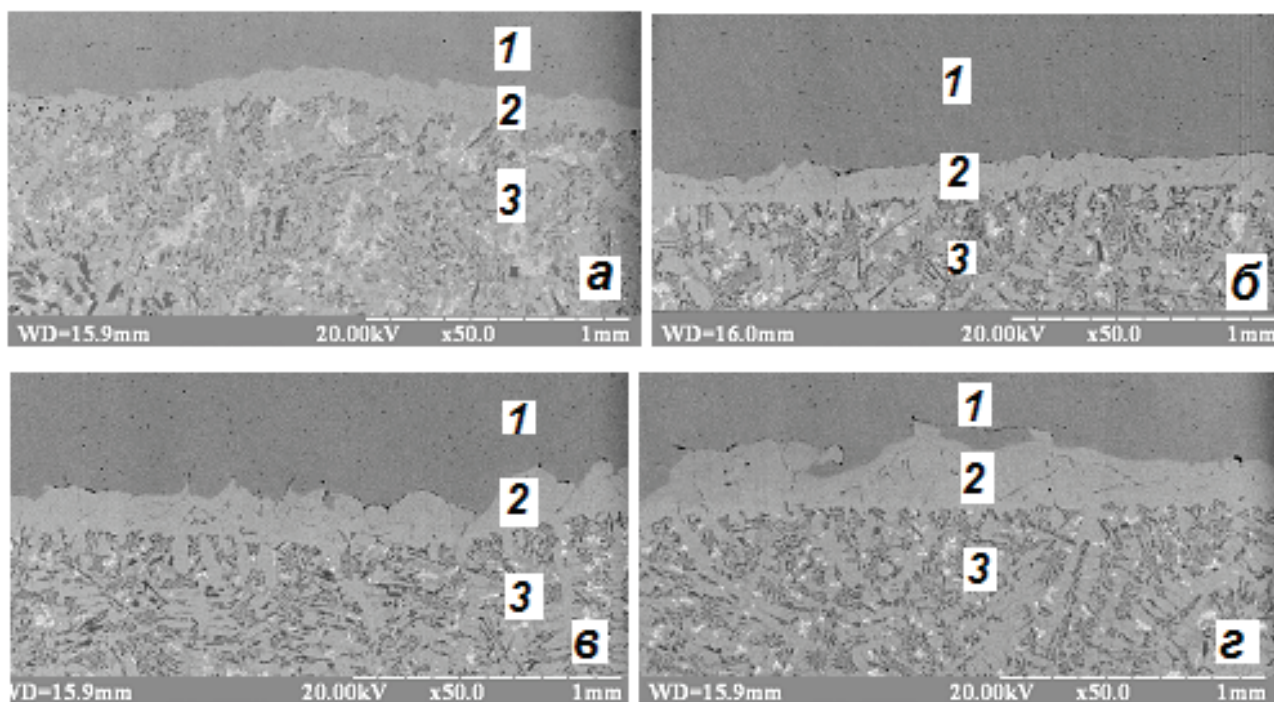
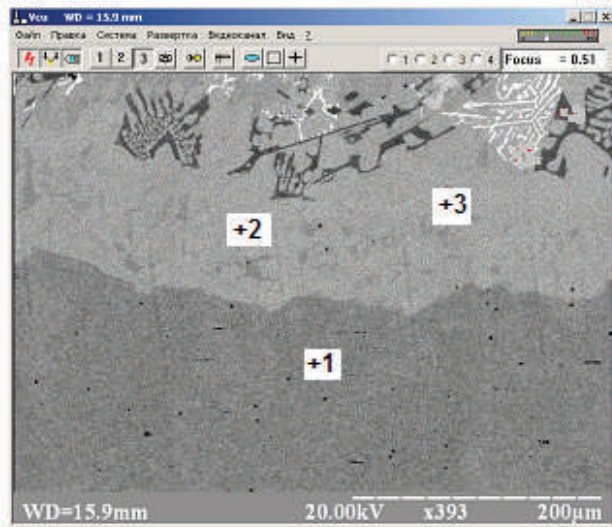


Рис. 2. Структура зони взаємодії СФС із залізом за температури 1200 °С за різного часу ізотермічної витримки (хв): 1 – залізо; 2 – зона взаємодії; 3 – самофлюсівний сплав; а – 10; б – 20; в – 30; г – 40



Информация			
Напряжение (кВ): 20.00			
Сумма %: 100.00			
Элемент	Инт.	С %	
Si	K 47	0.12	
Cr	K 2400	5.96	
Cr	L -1	0.00	
Fe	K 14408	85.71	
Fe	L -11	0.00	
Ni	K 564	5.59	
Ni	L -5	0.00	
Cu	K 38	0.46	
Cu	L -25	0.00	
Mo	L 413	2.17	
Mo	M 7	0.00	

Спектр 3

Информация			
Напряжение (кВ): 20.00			
Сумма %: 100.00			
Элемент	Инт.	С %	
Fe	K 21307	100.00	
Fe	L -3	0.00	

Спектр 1

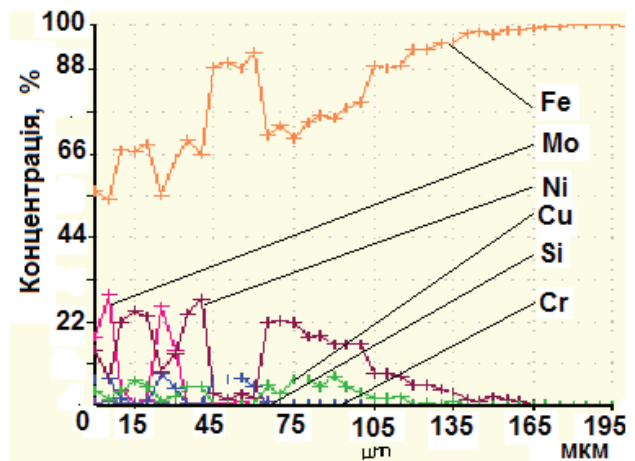
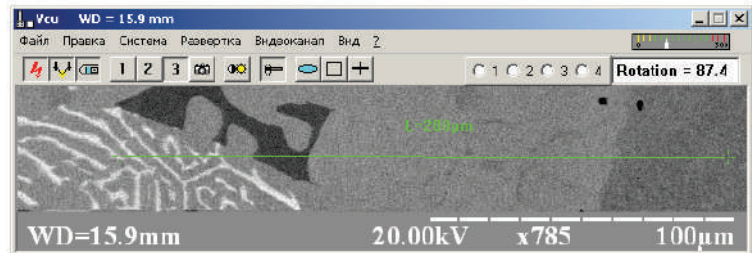


Рис. 4. Розподіл елементів на межі залізо – СФС після їх взаємодії за температури 1100 °С і часу ізотермічної витримки 10 хв

Информация			
Напряжение (кВ): 20.00			
Сумма %: 100.00			
Элемент	Инт.	С %	
Si	K 262	0.56	
Cr	K 458	0.29	
Cr	L -7	0.00	
Fe	K 16020	82.06	
Fe	L -10	0.00	
Ni	K 1569	13.69	
Ni	L -5	0.00	
Cu	K 319	3.40	
Cu	L -25	0.00	

Спектр 2

Як видно з рисунків, розплав СФС і залізо взаємодіють, і в межах контакту між ними утворюються нові фазові складники-прошарки,

ширина яких змінюється залежно від температури та часу ізотермічної витримки. У межах перехідної зони змінюється вміст елементів, а поблизу перехідних зон у СФС суттєво змінюється вміст хрому та нікелю. Водночас у залізі їх вміст, а також вміст кремнію та міді, збільшується. Останнє свідчить про те, що елементи, що входять до складу СФС, дифундують у залізо. Збільшення часу та температури взаємодії сприяє утворенню ширших проміжних прошарків на межі залізо – СФС із боку заліза, що узгоджується з наявними уявленнями про те, що розчинність високолегованих сплавів, до яких можна зарахувати СФС, більша, ніж у низьколегованих сплавів.

Наведені на рис. 3–4 мікроструктури свідчать про утворення широкої зони взаємодії, що пов'язано зі здатністю матеріалів взаємодіяти між собою, яка в свою чергу залежить від хімічної активності компонентів, і яку кількісно можна оцінити за вільною енергією хімічної реакції. А якщо реагуючі речовини перебувають не в чистому вигляді, а в розчиненому стані, то їх здатність взаємодіяти чи хімічна спорідненість додатково визначається їх концентрацією в розчині. Так, розчинність металів і сплавів у сплавах зменшується зі збільшенням ступеня легування останніх.

Одним із найважливіших факторів, що визначають процес взаємодії між сплавами на основі заліза з різним вмістом легуючих елементів, є дифузія. Як відомо, дифузійні процеси відбуваються тим швидше, чим сильніше природа елемента, що дифундує, відрізняється від природи розчинника. Це можна пояснити тим, що спотворення силових полів у кристалі поблизу атома розчиненого елемента тим сильніше, чим більше природа елемента відрізняється від природи розчинника. Відбувається вирівнювання концентрації між більш легованим СФС і залізом, у якому в початковому стані немає легувальних елементів.

Як зазначено вище, під час взаємодії СФС на основі заліза з залізом і його сплавами змінюється мікроструктура з утворенням перехідного шару. Самофлюсівний сплав на основі заліза має

евтектичну структуру [14], яка складається переважно з аустеніту, легованого нікелем, хромом, вуглецем, кремнієм, і карбоборидів типу MeB_2 , які можна ідентифікувати як $(Fe, Ni, Cr, Si)_2(B, C)$.

Під час взаємодії евтектичні структури розчиняються з утворенням світлої феритної фази, що є твердим розчином заліза і легуючих елементів із СФС, – це узгоджується з уявленнями І.І. Корнілова, який визначив сприятливу зону розчинності у межах 15 % (тобто приблизно 0,38 E) відхилень від атомного діаметра α -заліза. Отже, у фериті розчиняються кремній, нікель, хром, що підтверджується хімічним аналізом перехідного шару та результатами металографічних досліджень мікроструктур.

Легуючі елементи, розчиняючись у фериті, змінюють його властивості: підвищують корозійну стійкість, твердість, межу міцності та текучості, тобто опір пластичній деформації, а також опір в'язкому руйнуванню.

Виходячи з результатів дослідження процесів змочування та взаємодії розплаву СФС із залізом та його сплавами можна зробити висновок, що СФС можна застосовувати як джерело легувальних елементів при отриманні порошкових виробів на основі заліза. Водночас результати досліджень показують, що оптимальні технологічні режими (температура та час спікання) повинні бути в межах 1200–1250 °C за ізотермічної витримки понад 40 хв. У цьому разі відбувається повне змочування розплавом СФС заліза та суттєва дифузія легувальних елементів із СФС у залізо. Час витримки може змінюватись залежно від виду та вмісту легувальних елементів.

Як показано в роботах [15], [16], одним із перспективних матеріалів для отримання порошкових матеріалів на основі заліза для виробів спеціального призначення є сплав заліза легованого нікелем, молібденом, тому вивчались умови отримання порошкових матеріалів із заліза, легованого нікелем, хромом, вуглецем, міддю. Як вихідні матеріали використовували порошки карбонільного заліза, а як джерело легувальних елементів – порошок СФС. Їхні характеристики наведено в табл. 1.

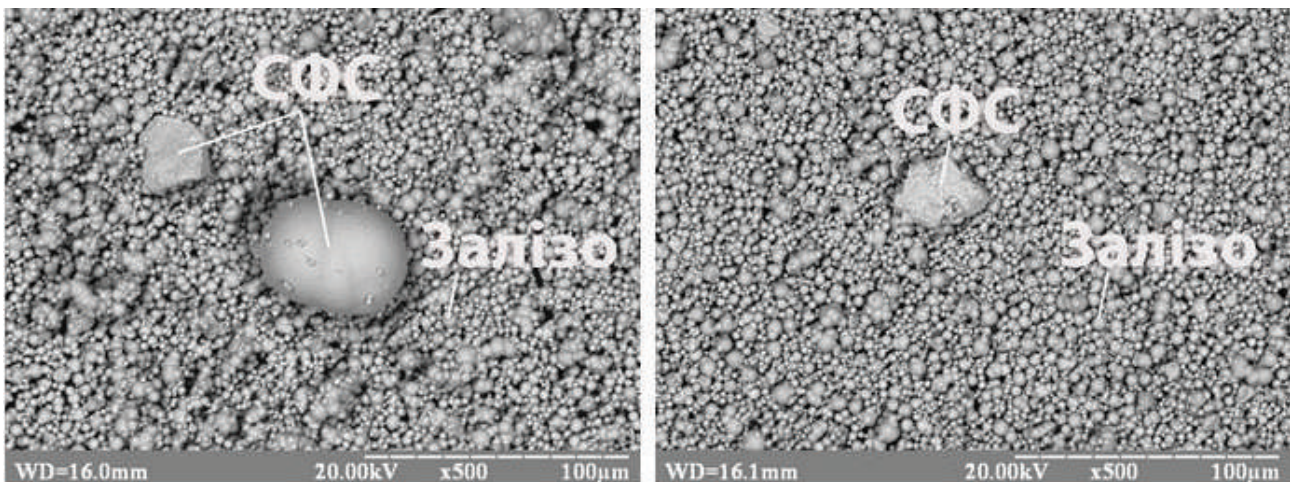
Таблиця 1. Характеристики вихідних порошків

Порошок	Хімічний склад, мас. %								Середній розмір частинок, мкм
	Fe	Si	B	Cr	Cu	P	C	Ni	
СФС	< 5	2,2	2,9	8	< 6	0,6	0,85	74,75	–
Залізо	99,9	0,05	–	–	–	–	0,05	–	< 15
Сплав Fe + 5 % СФС	95	0,11	0,145	0,4	0,3	0,03	0,042	3,723	–

Як видно з таблиці, розмір частинок порошку СФС значно перевищує розмір порошку заліза. Останнє є недоліком таких матеріалів, оскільки під час змішування площа контакту між ними мала, тож процеси гомогенізації матеріалу під час спікання уповільнюються. З урахуванням цього вивчався вплив розміру частинок вихідних порошків на процеси отримання матеріалів і, як наслідок, його кінцеві властивості. Збільшення контактної поверхні між частинками проводили в процесі змішування вихідної шихти з одночасним її подрібненням у планетарному млині. За цих обставин передбачали, що подрібнюватись буде насамперед порошок СФС, який має високу твердість і певну крихкість, що відповідає вимогам подрібнення матеріалів у планетарному млині. Вивчався вплив часу розмелювання на характеристики порошкових сумішей і виробів із них. Час розмелювання складав 0,5, 1, 2, 3 та 4 хв. Досліджено морфологію отримуваних порошків. Встановлено (рис. 5), що зі збільшенням часу розмелювання розмір частинок порошку заліза практично не змінюється, а частинок СФС – зменшується і за часу 4 хв стає практично сумірним із частинками порошку заліза (рис. 5, б). В отримуваних сумішах, особливо за розмелювання протягом 1 хв є частинки СФС із розміром до 50 мкм.

Досліджено процес ущільнення порошкових сумішей, отриманих за різного часу змішування–розмелювання. Пресування проведено в циліндричних пресформах діаметром 10 мм

за тиску пресування 400, 500, 600 і 700 МПа. Встановлено, що за всіх тисків пресування порошки вихідних сумішей мають низьку формуваність, яка зменшується для сумішей зі збільшення часу розмелювання–змішування, особливо за часу 3 та 4 хв. Такий характер ущільнення узгоджується з сучасними уявленнями про процеси ущільнення порошкових матеріалів. Відомо, що формуваність порошків за інших рівних умов залежить від якості контактів між частинками порошків, підпорядкованій площі контактів, процесам схоплювання, зварювання та ін. [11]. Найповніше ці процеси помітні під час ущільнення пластичних матеріалів. За наявності твердих і крихких матеріалів площа контакту між частинками погіршуватиметься (зменшуватиметься), бо в цьому разі немає пластичної деформації та відбувається лише точковий контакт. У нашому випадку зниження формуваності пресовок із порошкових сумішей, отриманих за часу розмелювання 3–4 хв, можна пояснити збільшенням питомої поверхні порошків СФС через зменшення їх розміру з підвищенням часу розмелювання. (рис. 6). Збільшення питомої поверхні порошків СФС запобігає утворенню якісного контакту між частинками порошку заліза та, як наслідок, зменшує формуваність. Тож для пресування порошкових сумішей залізо–СФС для підвищення їх формуваності необхідно використовувати зв'язки, які підвищують міцність контактів між частинками.



а

б

Рис. 5. Морфологія частинок порошкових сумішей Fe–СФС після розмелювання–змішування протягом різного часу; а – ... хв, б – 4 хв

Наявні зв'язки з поглядом їх технологічності та можливого внесення домішок у матеріал проаналізовано в статтях [6], [7]. Нами як зв'язку використано 2-відсотковий розчин полівінілового спирту у воді, який уводили в суміш вихідних порошків у кількості 1 об.%. Вивчення процесу пресування таких сумішей (рис. 6) показало, що вони мають задовільну формуваність. Щільність пресовок залежить від тиску пресування та часу розмелювання—змішування вихідних матеріалів. Отримані результати узгоджуються з сучасними уявленнями про вплив тиску, складу та властивостей вихідних порошкових сумішей. Оптимальним тиском пресування слід вважати 600 МПа.

Слід зауважити, що щільність пресовок, отриманих за оптимального тиску, дещо зменшується зі збільшенням часу розмелювання (рис. 6, 1). Останнє можна пояснити тим, що в цьому разі відбувається нагартування матеріалу та водночас зменшення ступеня пластичної деформації, яка переважно лежить в основі ущільнення досліджуваного матеріалу. Іншою причиною зменшення ступеня ущільнення може бути блокування пластичної деформації в зоні контакту між частинками заліза дисперсними частинками твердого СФС, кількість яких збільшується зі збільшенням часу розмелювання.

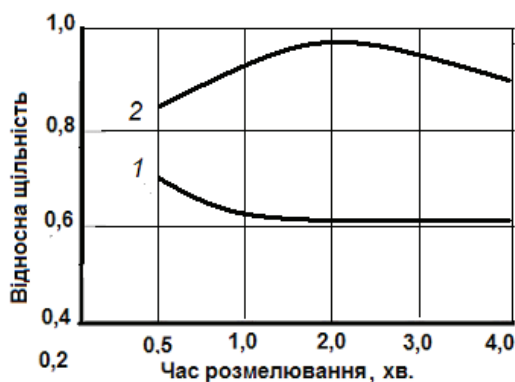


Рис. 6. Залежність відносної щільності від часу розмелювання: 1— після пресування; 2 — після спікання

Вивчалась також структура та деякі властивості матеріалів, отриманих з вихідних шихт, одержаних за різних умов. Зразки пресували за встановленим у роботі оптимальним тиском, а потім спікали в печі за температури 1250 °С із витримкою 80 хв задля підвищення щільності та формування заданих властивостей. Швидкість нагрівання становила 5 °С/хв. Такі умови, як було зазначено вище, найповніше сприяють отриманню щільних виробів і проход-

женню процесів сплавоутворення та гомогенізації матеріалів на основі заліза з використанням СФС як джерела легувальних елементів. Зразки, спечені за таких умов, мають щільність, яка змінюється залежно від часу розмелювання—змішування вихідних порошків у межах 5–10 %. Спостерігалась екстремальна залежність щільності матеріалу після спікання від часу розмелювання вихідної шихти (рис. 6, 2).

Характер цієї залежності можна пояснити таким чином. Відносно низька щільність матеріалу, виготовленого з шихти, отриманої змішуванням—розмелюванням протягом 0,5 хв, може бути зумовлена тим, що в такій шихті є частинки СФС відносно великого розміру (див. рис. 5), і у випадку їх плавлення через досконалу змочуваність, як показано в роботі [13], утворюються закриті пори, що спричинюють уповільнення процесу спікання. Зі збільшення часу розмелювання ступінь подрібнення частинок СФС збільшується, а вірогідність утворення закритих пор зменшується, що й сприяє утворенню матеріалу з практично стовідсотковою щільністю, отриманого з шихти розмелюванням протягом 2–3 хв. Зменшення щільності матеріалів з шихти, отриманої розмелюванням—змішуванням за часу понад 3 хв, може бути зумовлена деяким окисненням частинок порошків заліза. У цьому випадку за спікання відбувається взаємодія кремнію і бору, що входять до складу СФС, із утворенням нерозчинного боросилікатного шлаку [2], який розміщується на межі частинок заліза і в такий спосіб блокує процеси масоперенесення, які лежать в основі спікання [7]. Унаслідок цього щільність матеріалу, спеченого за інших рівних умов, зменшується.

Викладені міркування узгоджуються з результатами вивчення мікроструктури отриманих матеріалів (рис. 7).

Дослідження структури матеріалів після спікання показало, що матеріал має в основному полі зору практично стовідсоткову щільність. Поряд із цим зустрічаються фазові включення (пори, СФС) (рис. 7, а), причиною яких, як було зазначено вище, є наявність в шихті частинок СФС відносно великого розміру. Зі збільшенням часу змішування—розмелювання розмір частинок СФС зменшується, тож зменшується і вміст додаткових фазових включень, кількість яких у структурі матеріалу також зменшується (рис. 7, б, в), а за часу розмелювання 3 хв їх практично немає (рис. 7, г). Матеріали з такою структурою мають максимальну щільність після спікання.

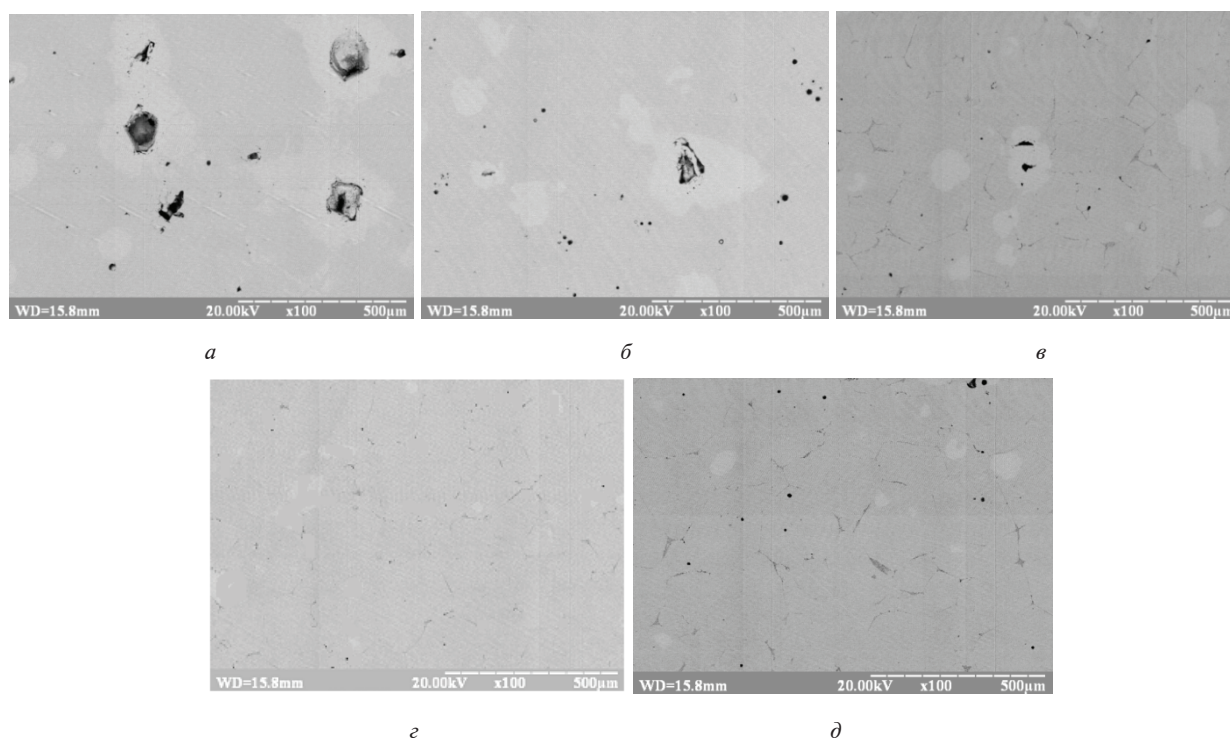
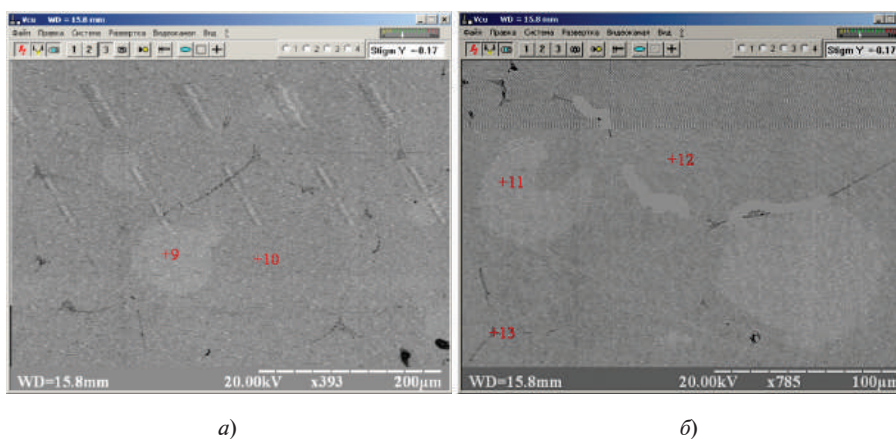


Рис. 7. Мікроструктура спеченого матеріалу з використанням шихти, отриманої змішуванням–розмелюванням порошків заліза та СФС протягом різного часу (хв): *а* – 0,5; *б* – 1; *в* – 2; *г* – 3; *д* – 4

У структурі матеріалів, отриманих із шихти змішуванням–розмелюванням протягом 4 хв теж з'являється фазовий складник на межі зерен заліза (рис. 8, *д*). Він може бути ідентифікований як боросилікатний шлак, який виникає за причин, викладених вище. Наявність його на межі фазових складових також може спричинити вповільнення сплавоутворення (гомогенізації).

Вивчення структури отриманих матеріалів свідчить, що за спікання СФС взаємодіє із залізом, що узгоджується з вищенаведеними отриманими в роботі результатами. Відбувається й пере-

важна дифузія складників СФС у залізо (рис. 8). За досліджуваних умов спікання розподіл елементів у кінцевому матеріалі дещо нерівномірний і залежить від гранулометричного складу вихідної шихти, який залежить (див. рис. 8) від вихідних матеріалів й умов змішування–розмелювання. Наявність легувальних елементів в отриманому матеріалі підтверджується результатами якісного аналізу розподілу елементів (рис. 9), який доводить наявність у ньому заданих елементів і рівномірність їх розподілу у сплаві.



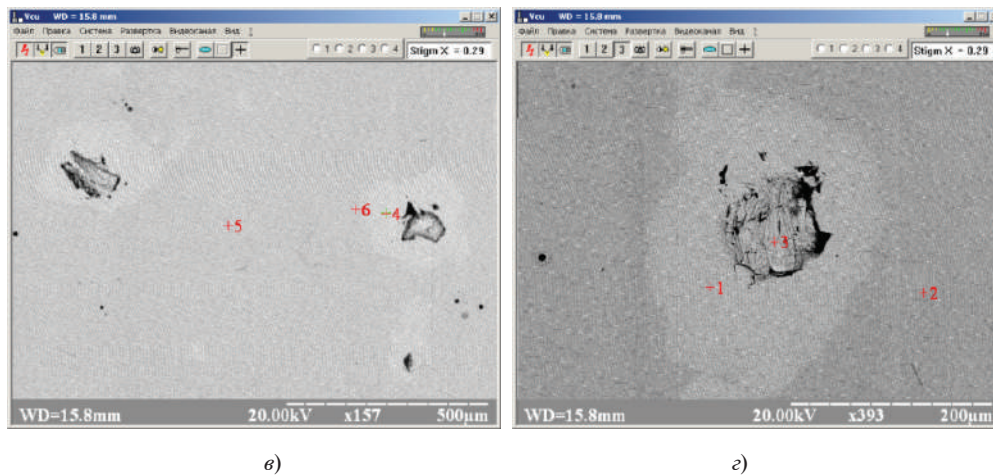


Рис. 8. Розподіл елементів у структурі матеріалу Fe–СФС, отриманого з використанням шихти, одержаної за різного часу змішування–розмелювання.

а) Час змішування 0,5 хв; 1 – Fe – 69,63; Ni – 26,73; Cr – 2,3; Si – 1,34; 2 – Fe – 97,73; Ni – 1,76; Cr – 0,32; Si – 0,19; 3 – Fe – 76,57; Ni – 17,48; Cr – 1,18; Si – 0,19.

б) Час змішування – 1 хв. 4 – Fe – 75,61; Ni – 21,36; Cr – 2,3; Si – 1,11; 5 – Fe – 98,47; Ni – 0,95; Cr – 0,34; Si – 0,24; 6 – Fe – 95,51; Ni – 3,62; Cr – 0,44; Si – 0,44

в) Час змішування – 3,0 хв. 11 – Fe – 91,91; Ni – 6,02; Cr – 1,41; Si – 0,67; 12 – Fe – 95,74; Ni – 3,01; Cr – 0,70; Si – 0,55; 13 – Fe – 96,93; Ni – 1,40; Cr – 1,47; Si – 0,20

г) Час змішування – 4,0 хв. 9 – Fe – 89,89; Ni – 9,10; Cr – 0,54; Si – 0,47; 10 – Fe – 96,45; Ni – 3,02; Cr – 0,11; Si – 0,42.

Для з'ясування впливу ступеня легування на властивості матеріалу вивзначалась твердість зразків після спікання залежно від умов отримання.

Залежність твердості від умов отримання вихідної шихти наведена на рис. 10.

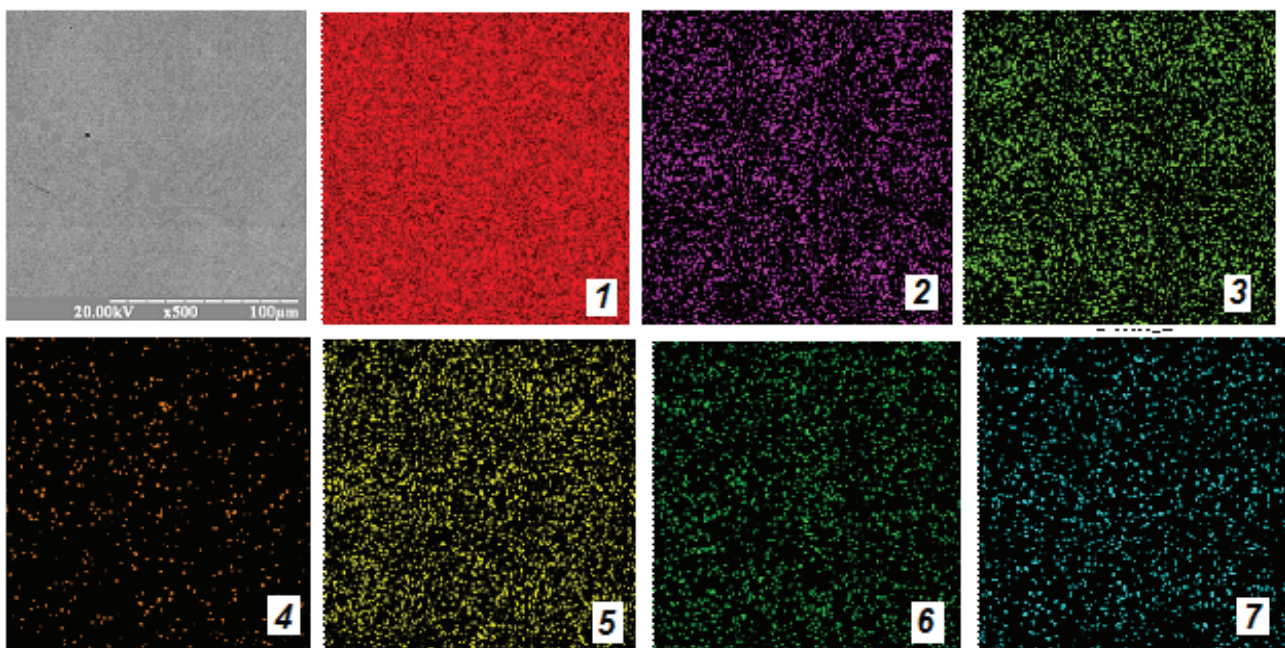


Рис. 9. Розподіл елементів у матеріалі Fe–СФС: 1 – Fe; 2 – Ni; 3 – Cr; 4 – Si; 5 – Cu; 6 – C; 7 – B

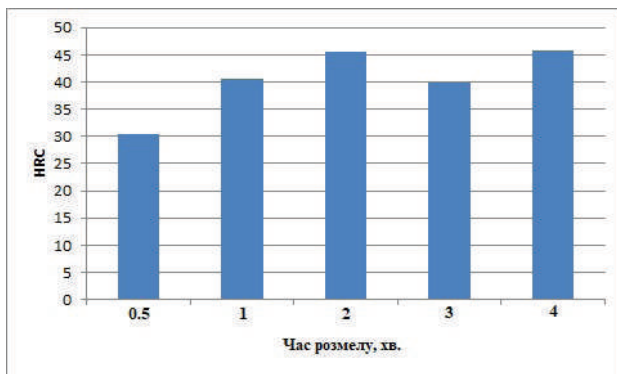


Рис. 10. Залежність твердості матеріалу від часу розмелювання–змішування вихідної суміші AN9–Fe під час його отримання

Як видно з рис. 10, на твердість матеріалу впливають умови отримання вихідної шихти змішуванням–розмелюванням. Отримані значення твердості більші, ніж у чистого заліза, що свідчить про проходження процесів легування під час спікання. При цьому має місце екстремальна залежність твердості кінцевих матеріалів від часу приготування шихти. Збільшення часу від 0,5 до 3 хв сприяє підвищенню твердості, і надалі вона дещо зменшується. Такий характер залежності можна пояснити, виходячи з того, що твердість залежить від хімічного та фазового складу, структури матеріалу, його гомогенності та ін. [12]. Якщо припустити, що хімічний склад матеріалу зразків, отриманих із різної шихти, однаковий, то залежність твердості від часу приготування вихідної шихти слід пов'язувати зі ступенем легування та гомогенністю отриманих матеріалів, які, як було встановлено вище, збільшуються зі збільшенням часу розмелювання–змішування під час отримання вихідної шихти. Деяке зменшення твердості зі збільшенням часу отримання вихідної шихти, більшому

за 3 хв, слід пов'язувати з появою у структурі матеріалу на межах зерен боросилікатного шлаку за вищевказаних причин. Він з одного боку є твердим, але дуже крихким матеріалом, а з іншого – погіршує міцність зв'язку між фазовими складниками та ступінь взаємодії між ними. Усі ці процеси призводять до зменшення твердості.

Висновки

Вивчено процеси сплавоутворення під час отримання порошкового матеріалу на основі заліза з умістом легувальних елементів нікелю, хрому, міді, силіцію, вуглецю, бору та інших у межах 0,5–5 мас.% із використанням як джерела легувальних елементів СФС, які є високолегованими сплавами та вміщують усі ці елементи. Показано, що перспективним методом отримання таких матеріалів є готування вихідної шихти з вихідних порошоків заліза та СФС, її пресування та спікання за наявності рідкої фази. Досліджено процеси змочування розплавами СФС заліза та взаємодії між ними. Встановлено оптимальні режими спікання: температура понад 1200 °С за часу ізотермічної витримки більш ніж 80 хв, які повинні забезпечувати практично стовідсоткову щільність матеріалу та його гомогенність.

Встановлено, що під час спікання відбувається легування заліза елементами, які входять до складу СФС, завдяки уніполярній дифузії. Показано, що на гомогенність кінцевого матеріалу впливає якість змішування–розмелювання. Зменшення розміру частинок порошоків вихідних матеріалів сприяє отриманню більшого ступня легування, гомогенності матеріалу та підвищенню його механічних характеристик.

У подальшому бажано встановити оптимальні режими спікання для отримання більш гомогенних сплавів.

References

- [1] A. M. Stepanchuk, *Theoretical and technological foundations for the production of powders of metals, alloys and refractory compounds*. Kyiv, Ukraine: NTUU “KPI”, 2007.
- [2] A. M. Stepanchuk and I. I. Bilyk, *Materials for spraying coatings. Textbook*. Kyiv, Ukraine: Tsentr uchbovoi literatury, 2016.
- [3] I. D. Radomyselskiy and S. G. Napara-Volgina, *Obtaining alloyed powders by diffusion method and their use*. Kyiv, USSR: Naukova dumka, 1988.
- [4] E. S. Nichiporenko, Yu. I. Naida, and A. B. Medvedovsky, *Sprayed metal powders*. Kyiv, USSR: Naukova dumka, 1980.
- [5] Yu. F. Ternovoy, G.A. Baglyuk, and S. S. Kudievsky, *Theoretical foundations of processes of spraying of metallic melts*. Zaporozhye, Ukraine: Vydavnytstvo Zaporizkoi derzhavnoi inzhenernoi akademii, 2008.
- [6] A. M. Stepanchuk, *Theory and technology of pressing powder materials. Textbook*. Kyiv, Ukraine: Tsentr uchbovoi literatury, 2016.
- [7] A. N. Stepanchuk, I. I. Bilyk, and P. A. Boyko, *Powder metallurgy technology*. Kyiv, USSR: Vyshcha shkola, 1989.
- [8] A. N. Stepanchuk, A.A. Nechiporenko and N.S. Vdovichenko, “Obtaining and properties of self-fluxing alloys based on iron”, *Suchasni dosiahnennia v haluzi tekhniki ta zastosuvannia hazotermichnykh ta vakuumnykh pokryttiv*, pp. 78 – 84, 1991.

- [9] V. N. Eremenko, Ed., *Wettability and surface properties of melts and solids*. Kyiv, USSR: Naukova dumka, 1972.
- [10] S. I. Popil, *Surface phenomena in melts*. Moscow, Russia: Metalurhiia, 1994.
- [11] A. M. Stepanchuk, O. A. Demidenko and L. M. Lopushanska, "Wetting and interaction of melts of self-fluxing alloys of iron and its alloys", In *Proc. of the international scientific and technical conference on Materials for work in extreme conditions – 6'03*, 2016, pp. 106–109.
- [12] V. N. Zaplatin, Yu. I. Sapozhnikov and A.V. Dubov, *Fundamentals of materials science (metalworking): a textbook for students of middle prof. education institutions, 8th ed.* Moscow, Russia: Vydavnychiy tsentr "Akademiia", 2017.
- [13] O.A. Demydenko and A.M. Stepanchuk, "Influence of composition and technological factors of obtaining compositional materials of iron-self-flusive alloy on physico-mechanical characteristics", *Materials Science and Technology. European Science*, vol. 2, pp. 126–131, 2020.
- [14] A. M. Stepanchuk, O. I. Goncharuk, I.V. Goltvenko and M.A. Sklyarov, "Structure and some properties of self-fluxing alloys based on iron and coatings from them", "*Naukovi notatky*" LGTU, no. 13, pp. 319 – 334, 2003.
- [15] S. V. Zavadiuk, P. I. Loboda, T. O. Soloviova, I. Iu. Trosnikova and O. P. Karasevska, "Fracture Features of Low-Alloy Steel Produced by Metal Injection Molding", *Powder Metallurgy and Metal Ceramics*, vol. 59, pp. 641–650, 2021. doi: 10.1007/s11106-021-00199-0
- [16] S.V. Zavadiuk, P.I. Loboda, T.O. Soloviova, I.Iu. Trosnikova and O.P. Karasevska, "Optimization of the Sintering Parameters for Materials Manufactured by Powder Injection Molding", *Powder Metallurgy and Metal Ceramics*, vol. 59, pp. 22–28, 2020. doi: 10.1007/s11106-020-00134-9.
- [17] V.V. Skorokhod and S.M. Solonin. *Physical and metallurgical foundations of sintering powders*. Moscow, USSR: Metalurhiia, 1984.

D.O. Remizov, A.M. Stepanchuk, T.O. Soloviova, P.I. Loboda, S.V. Zavadiuk

THE USE OF SELF-FLUXING ALLOYS AS A SOURCE OF ALLOYING ELEMENTS IN PRODUCING POWDER PRODUCTS OF CONSTRUCTION PURPOSE FROM LOW-ALLOYED IRON

Background. One of the problems in the production of powder structural materials based on iron alloyed with a small amount of alloying elements (0.5 - 5.0 wt.%) is obtaining a homogeneous alloy in the final material. The distribution uniformity of the alloying elements largely determines the final properties of the material. The use of a mixture of powders constituting the alloy to obtain such alloys, in many cases, does not solve the problem of obtaining homogeneous alloys. Promising in this respect is the use of alloy powders containing alloying elements as sources of alloying elements.

Objective. The purpose of the paper is to study the formation regularities of the material's structure and properties based on iron alloys obtained using self-fluxing alloys (SFA) as a source of alloying elements.

Methods. To solve the problem posed in the work, the methods of mixing-grinding in planetary mills, the study of the processes of wetting with SFA iron melts and the interaction between them by the method of a lying drop, scanning microscopy for studying the structure of materials and the distribution of elements in them, measuring mechanical properties (hardness) were used.

Results. The principles of determining the optimal conditions for obtaining materials - grinding-mixing time (4 min.), compacting pressure of the original workpieces (600 MPa), sintering temperature (1200 °C), sintering time (more than 80 min.) were established. Their use contributes to the production of products with almost one hundred percent material density and a high degree of homogeneity.

Conclusions. As a result of the work carried out, it was found that the use of SFA as alloying elements sources and the preparation of the initial mixture by grinding-mixing methods in a planetary mill promotes the production of materials with a high degree of homogeneity due to the intensification of homogenization and sintering processes and, as a consequence, an increase in their mechanical characteristics.

Keywords: iron, alloying elements, grinding, pressing, sintering, material structure, hardness.

Рекомендована Радою
інституту матеріалознавства та зварювання ім. Є.О. Патона
КПІ ім. Ігоря Сікорського

Надійшла до редакції
11 листопада 2021 року

Прийнята до публікації
14 лютого 2022 року

DOI: 10.20535/kpissn.2021.4.250663

УДК 628.165

Серета В.В.^{*}, Соломаха А.С., Припула Н.О., Заболотний О.А.

КПІ ім. Ігоря Сікорського, Київ, Україна

^{*}corresponding author: volodya.81.vs@gmail.com

ТЕРМОДИНАМІЧНИЙ АНАЛІЗ ТЕРМІЧНОЇ ОПРІСНЮВАЛЬНОЇ УСТАНОВКИ З ЦИКЛОМ ЗВОЛОЖЕННЯ-ОСУШЕННЯ ПОВІТРЯ

Проблематика. Сьогодні опріснення морської води забезпечує видобуток світової прісної води у кількості 24,5 млн. м³ на добу. Одним із перспективних способів опріснення є термічна технологія зволоження–осушення повітря. Незважаючи на суттєві переваги цієї технології, ключовими проблемами її використання є високі енергетичні потреби.

Мета дослідження. Визначити оптимальні термодинамічні параметри робочих тіл (повітря і води) в конструктивних елементах опріснювальної установки з циклом зволоження–осушення повітря для отримання максимальної продуктивності за мінімального підведення теплоти.

Методика реалізації. Математичне моделювання циклу зволоження–осушення повітря за різних режимних параметрів на основі створеної термодинамічної моделі з масовими та енергетичними балансами кожного елемента опріснювальної установки.

Результати дослідження. Встановлено, що енергетична ефективність (gain out put ratio, GOR) установки збільшується зі збільшенням температури солоної води на виході з осушувача. Максимальні значення GOR та коефіцієнта відновлення води (recovery ratio, RR) спостерігаються за максимального перепаду температур (максимального нагрівання) солоної води в осушувачі. Показано, що зі збільшенням температури солоної води на вході до зволожувача GOR та RR також збільшуються. З'ясовано, що GOR не залежить від температури та вологості повітря на вході до зволожувача.

Висновки. Максимальні значення GOR та RR для циклу з додатковим підігріванням води становлять 3,75 та 6% відповідно. Для отримання таких значень температура солоної води на вході в установку (осушувач) має дорівнювати 10 °С, а температура солоної води на вході до зволожувача – 70 °С за будь-яких температур та вологості повітря. Застосування відновлювальної енергії для підігрівання солоної води у нагрівнику дасть змогу значно підвищити GOR.

Ключові слова: енергетична ефективність, зволоження–осушення повітря, математична модель, термічне опріснення, термодинамічний аналіз.

Вступ

Попит на чисту питну воду зростає як в Україні, так і в усьому світі через збільшення чисельності населення, урбанізацію, індустріалізацію, сільськогосподарську діяльність і соціально-економічний розвиток.

Споживання прісної води зросло в шість разів між 1900 і 1995 роками, що більш ніж удвічі порівняно зі зростанням населення. Нині із близько 6,8 млрд. населення світу приблизно одна третина

проживає у країнах, що мають проблеми з прісною водою [1]. У 2050 році ця цифра може зрости до 5,7 млрд. [2]. Нині 3,5 млн. людей щорічно гинуть у результаті неналежного водопостачання та каналізації, що посилює роль води як важливого глобального ресурсу [3]. З огляду на це потреба в прісній воді стоїть на одному з перших місць у міжнародному порядку денному критичних проблем (поряд зі зміною клімату).

Приблизно половина європейських країн уже зараз стикається з нестачею води, що вимагає

Рекомендуємо цитувати цю статтю так: В.В. Серета, А.С. Соломаха, Н.О. Припула, О.А. Заболотний, «Модель термічної опріснювальної установки з циклом зволоження-осушення повітря», *Наукові вісті КПІ*, № 4, с. 69–76, 2021. doi: 10.20535/kpissn.2021.4.250663

Please cite this article as: V.V. Sereda, A.S. Solomakha, N.O. Prytula, O.A. "Thermodynamic analysis of thermal desalination system with humidification–dehumidification cycle", *KPI Science News*, no. 4, pp. 69–76, 2021. doi: 10.20535/kpissn.2021.4.250663.

комплексних управлінських зусиль, щоб збалансувати попит і пропозицію. Великі території у Сполучених Штатах використовують більше води, ніж можна поповнити природним шляхом. В Індії є райони, які стикаються з постійним дефіцитом води, і велика кількість сіл має проблеми із засоленням води. В Ізраїлі забір води також перевищує теоретичну кількість, незважаючи на багато винайдених технологій збереження води. Ситуація також серйозна в Китаї, де у 550 із 600 найбільших міст країни не вистачає води. Зростаючий попит на воду потенційно може призвести до зниження доступності для побутового, сільськогосподарського та промислового використання. Крім того, китайські річки настільки забруднені важкими металами, що їх не можна використовувати для зрошення. Це неминуче призводить до втрат виробництва, як промислового, так і сільськогосподарського. Єгипет також змушений імпортувати більше від половини продовольства, оскільки у нього не вистачає води для вирощування всередині країни. Україна вже за декілька десятиріч може зіткнутися з дефіцитом якісної питної води [4]. Механізм, що призводить до дефіциту води у країнах, що розвиваються, часто дуже складний, діють взаємопов'язані чинники, такі як стрімкий приріст населення, брак інфраструктури та обмеженість природних водних ресурсів [1].

Другою важливою сучасною проблемою є виявлення та видалення нових хімічних забруднювачів, таких як гідрофільні органічні сполуки, побічні продукти дезінфекції, фармацевтичні сполуки, а також багато різних іонів, що походять з електронних виробів, які часто потрапляють на звалища, забруднюючи таким чином землю і воду.

У результаті такої ситуації повторне використання та опріснення води стали ключовими чинниками для підтримки сталого розвитку. Доказом цього є стабільне зростання встановленої потужності опріснювальних установок у всьому світі. Забезпечення сталого розвитку світового водопостачання вимагає сьогодні розроблення та впровадження передових та екологічно прийнятних процесів, які б дали змогу зберегти воду та зменшити її споживання. Одним із таких процесів є опріснення води за допомогою зволоження-осушення повітря (humidification dehumidification, HDH).

Для опріснення води використовують термічні або мембранні технології [5]. Мембранні технології нині домінують на ринку опріснення. Дев'яносто відсотків контрактів на будівниц-

тво опріснювальних установок, укладених із 2010 року, використовують мембрани. Мембранні технології набули такого поширення через свої незаперечні переваги: можливість організації безперервного опріснення; висока швидкість опріснення; високі показники відновлення води, видалення солей, низьке споживання енергії; мінімальні вимоги до попереднього очищення води, тривалий строк експлуатації [6]. Разом із цим, мембранам притаманні й суттєві недоліки, зокрема: пестициди, гербіциди та хлор молекулярно менші за воду і можуть проходити через мембрану; ймовірне видалення природних мінералів із води; значні витрати води, яка проходить через систему; погіршення якості води з часом експлуатації; потреба в енергії високої якості (електрична енергія, високотемпературне тепло), що не дає змоги напряму комбінувати мембрани із сонячною або іншою альтернативною енергією [7].

Термічні процеси застосовують для великомасштабних проєктів, зосереджених на Близькому Сході [8]. Термічні технології передбачають підігрівання солоної води і збирання конденсованої пари (дистиляту) для отримання чистої води. Найбільш поширені термічні багатоступеневі установки миттєвого скипання та з регенерацією отриманої пари в наступних ступенях. Незважаючи на велику кількість переваг термічного опріснення, ці технології останнім часом рідко використовують через викиди забруднювальних речовин в атмосферу та значне споживання енергії. Однак термічне опріснення ще застосовують на установках великої продуктивності [9]. В останні роки значні зусилля науковці спрямували на дослідження термічних опріснювальних установок із застосуванням сонячної енергії, що дає змогу суттєво підвищити ефективність процесу дистиляції [10].

Технологія HDH є порівняно новим методом термічного знесолення, який використовує повітря як носій для опріснення солоної води. Процес HDH підходить для децентралізованого немасштабного виробництва прісної води. Такі системи є ефективними у виробництві прісної води для невеликих громад, особливо у віддалених місцях, які не підключені до електричної або водорозподільної мережі [11]. HDH-системи мають просту конструкцію, нескладні у виготовленні, не потребують складного технічного обслуговування і можуть працювати від низькотемпературних або відновних джерел енергії. Порівняно з мембранними технологіями, HDH-системи очищають воду з високими концентраціями

солей (фактично будь-якими); на виході отримуємо 100% дистилат (конденсат); не погіршується якість води із часом експлуатації [8].

Головним недоліком НДН-систем є високі енергетичні потреби. З термодинамічного та екологічного погляду доцільніше використовувати відпрацьоване тепло або відновлювані джерела енергії, ніж електричну енергію або природний газ. Хоча такі низько потенційні джерела доступні з мінімальними експлуатаційними витратами, треба визначити оптимальну теплову конструкцію НДН-системи для отримання максимальної швидкості виробництва води за мінімального підведення теплоти. У відкритій літературі бракує інформації щодо зв'язку термодинамічного аналізу з оптимальною інженерною конструкцією НДН-систем. Так, у наукових працях [12–14] подані математичні моделі циклу зволоження-осушення повітря з нагріванням води. Досліджено вплив швидкості потоку сирової води та розміру зволожувача та осушувача на швидкість виробництва прісної води. Однак теплових характеристик, які не залежать від швидкості виробництва води та розміру обладнання, не було подано.

Постановка задачі

Надати комплексний термодинамічний аналіз циклу зволоження-осушення повітря з підігріванням води та подати оптимальні умови його роботи.

Конструктивні особливості НДН-систем

Схему опреснювальної установки, яка працює за принципом зволоження-осушення повітря, показано на рис. 1. Повітря та солоня вода рухаються у розімкнених контурах. Солоня вода насосом подається до осушувача з температурою $t_{1в}$. Далі солоня вода циркулює по трубках осушувача, нагріваючись до температури $t_{2в}$ за рахунок відведення теплоти від насиченого повітря, яке осушується за рахунок охолодження у міжтрубному просторі осушувача.

Навколишнє повітря з температурою $t_{1п}$ і вологістю $d_{1п}$ вентилятором подається у зволожувач. У зволожувачі повітря нагрівається до температури $t_{2пн}$ зволожується за рахунок безпосереднього контакту з підігрітою у нагрівнику до температури t_3 всолоню водою. Гаряча вода розбризкується через форсунки для ефективного тепло- і масообміну. Графік зміни температур повітря та солоні води у зволожувачі показано на рис. 2, а.

Частина води випаровується в повітря, тоді як решта виділяється як солоний залишок у нижній частині зволожувача. На виході зі зволожувача вологість повітря може досягнути значення 100 %. Далі насичене повітря надходить до осушувача, де водяна пара конденсується з повітря для виробництва прісної води. Охолоджене повітря з температурою $t_{3п}$ видаляється в навколишнє середовище. Графік зміни температур повітря та солоні води в осушувачі показано на рис. 2, б.

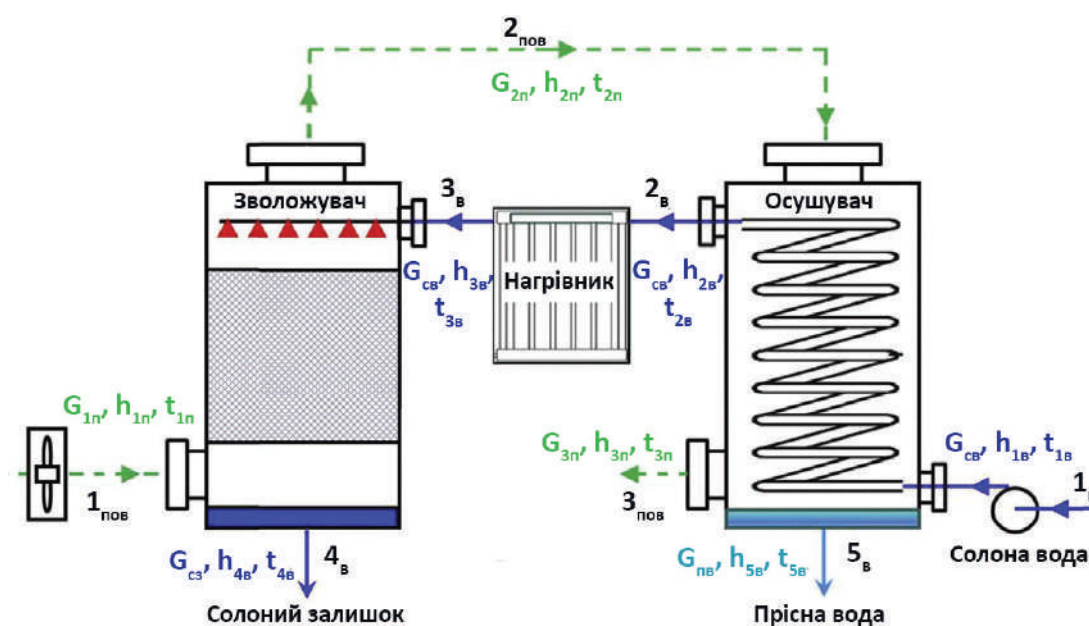


Рис. 1. Опреснювальна установка, яка працює за принципом зволоження-осушення повітря

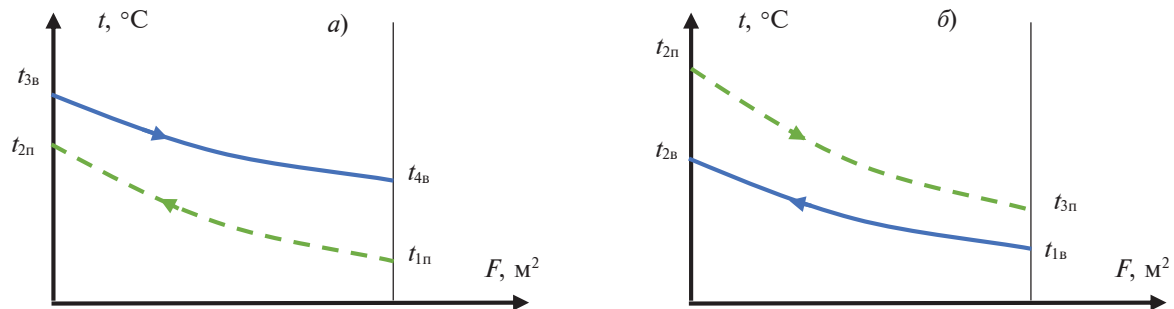


Рис 2. Перепад температур в установці: а) зволожувач; б) осушувач.

Розроблення математичної моделі

Для моделювання циклу, показаного на рис. 1, були зроблені такі припущення:

- цикл працює за стаціонарних умов;
- втрати теплоти від будь-якого обладнання установки не враховуються;
- потужність насоса та вентилятора є незначною порівняно з тепловою енергією, яка підводиться у нагрівнику;
- прісна вода виходить з осушувача з температурою, яка є середньою між температурою точки роси повітря на вході до осушувача й температурою повітря на виході з осушувача;
- відносна вологість повітря, що виходить зі зволожувача та осушувача, становить 100 %.

Модель, подана в цій роботі, є термодинамічною моделлю, де масові та енергетичні баланси застосовують до кожної зі складових циклу, дотримуючись першого закону термодинаміки [15]. Однак для математичного аналізу циклу та розрахунку його продуктивності треба знати деякі інші параметри. Це – температура води на вході до осушувача повітря $t_{1в}$ і температура води на вході до зволожувача повітря $t_{3в}$. Мінімальна температура води в циклі на вході в установку $t_{1в}$ може коливатися між 10–40 °C через сезонні зміни температури. Максимальна температура води в циклі на вході до зволожувача $t_{3в}$ має міститися в діапазоні 50–70 °C для недопущення утворення накипу. Ключові рівняння моделі мають такий вигляд:

1. Енергетичний та масовий баланси осушувача повітря:

$$G_{св} h_{1в} + G_{2п} h_{2п} = G_{св} h_{2в} + G_{пв} h_{5в} + G_{3п} h_{3п}; \quad (1)$$

$$G_{пв} = G_{2п} (d_{2п} - d_{3п}). \quad (2)$$

2. Ефективність осушувача визначають за такою формулою:

$$\varepsilon_{ос} = \max \left(\frac{h_{2п} - h_{3п}}{h_{2п} - h_{2п}^{ideal}}; \frac{h_{2в} - h_{1в}}{h_{2в}^{ideal} - h_{1в}} \right), \quad (3)$$

де $h_{3п}^{ideal}$ – ідеальна ентальпія вихідного повітря обчислюється за умови повного насичення повітря на виході, за $t_{3п} = t_{1в}$; $h_{2в}^{ideal}$ – ідеальна ентальпія солоні води на виході з осушувача, коли $t_{2в} = t_{2п}$.

3. Енергетичний та масовий баланси зволожувача повітря:

$$G_{1п} h_{1п} + G_{св} h_{3в} = G_{2п} h_{2п} + G_{с3} h_{4в}; \quad (4)$$

$$G_{с3} = G_{св} - G_{пв}. \quad (5)$$

4. Ефективність зволожувача визначаємо аналогічно формулі :

$$\varepsilon_{зв} = \max \left(\frac{h_{2п} - h_{1п}}{h_{2п}^{ideal} - h_{1п}}; \frac{h_{3в} - h_{4в}}{h_{3в} - h_{4в}^{ideal}} \right), \quad (6)$$

де $h_{2п}^{ideal}$ ідеальна ентальпія вихідного повітря за умови повного насичення, за $t_{2п} = t_{3в}$; $h_{4в}^{ideal}$ ідеальна ентальпія солоні залишку на виході зі зволожувача, коли $t_{4в} = t_{1п}$.

5. Енергетичний баланс нагрівника:

$$Q_{наг} = G_{св} (h_{3в} - h_{2в}). \quad (7)$$

Математичну модель, подану рівняннями (1)–(7), розв'язують за допомогою інженерного програмного забезпечення РТС Mathcad, використовуючи для визначення властивостей вологого повітря та солоні води пакет CoolProp [16]. Після вирішення рівнянь (1)–(7) розраховують такі параметри продуктивності:

1. Коефіцієнт енергетичної ефективності (gainoutputratio, GOR) – основний показник оцінювання роботи опріснювальної установки. Він показує, скільки теплової енергії споживається в процесі опріснення. Чисельно дорівнює відношенню витрати виробленої прісної води,

помноженої на приховану теплоту випаровування, до підведеної у циклі теплової енергії:

$$GOR = G_{\text{пв}} r / Q_{\text{наг}} \quad (8)$$

2. Коефіцієнт відновлення води (recovery ratio, RR) – відношення витрати прісної води, яка отримується в установці, до вхідної солоної води:

$$RR = G_{\text{пв}} / G_{\text{св}} \quad (9)$$

3. Коефіцієнт масової витрати (massflow-ratio, MR) – відношення масової витрати вхідної солоної води до масової витрати зовнішнього повітря:

$$MR = G_{\text{св}} / G_{\text{пн}} \quad (10)$$

Результати моделювання

Як бачимо на рис. 3, значення GOR збільшується зі збільшенням температури солоної води на виході з осушувача $t_{2в}$. Тобто чим більше теплоти буде відведено від насиченого повітря в осушувачі (див. формулу), тим більше сконденсується прісної води з повітря і тим ефективнішою буде робота установки. Значення температури $t_{2в}$ залежать як від перепаду температур солоної води в осушувачі $\Delta t_{св}$, так і від температури солоної води на вході в установку $t_{1в}$:

$$t_{2в} = t_{1в} + \Delta t_{св} \quad (11)$$

Також рис. 3 показує, що за однакових $t_{2в}$ значення GOR тим більше, чим менша $t_{1в}$. Це зумовлено тим, що зменшення $t_{1в}$ дає змогу зменшити температуру $t_{3п}$ і вологовміст $d_{3п}$ повітря на виході з осушувача (див. рис. 2 та 4). З формули (2) випливає, що охолодження повітря до меншої температури дає змогу сконденсувати більше свіжої води з повітря. Однак температуру $t_{1в}$ у цьому циклі регулювати немає змоги. Її значення залежить від сезонних змін і становить 10–40 °C. З огляду на це для забезпечення максимальних значень GOR та RR необхідно, щоб перепад температур солоної води в осушувачі $\Delta t_{св}$ був якомога більшим.

Розглянемо на скільки максимально доцільно збільшити $\Delta t_{св}$ і чи можливо досягти рівності температур $t_{2в}$ і $t_{3в}$, щоб позбутися нагрівника у цьому циклі. Зі збільшенням $\Delta t_{св}$ збільшується кількість сконденсованої прісної води з повітря, що відповідно до формули (2) призведе до збільшення витрати повітря в осушувачі (за постійних температур $t_{2п}$ та $t_{3п}$). Відповідно кількість

повітря, яка проходить через зволожувач, теж зросте, а значення коефіцієнта масової витрати MR зменшуватиметься (див. рис. 5). Щоб нагріти більшу кількість повітря у зволожувачі, треба підвести більше теплоти від солоної води. З рівняння теплового балансу зволожувача бачимо, що чим більше теплоти відводиться від солоної води, тим температура солоного залишку $t_{4в}$ буде меншою. Натомість розподіл температур у зволожувачі (рис. 2) показує, що температура $t_{4в}$ не може бути меншою за температуру зовнішнього повітря $t_{1п}$, тобто солоний залишок неможливо охолодити нижче від температури зовнішнього повітря.

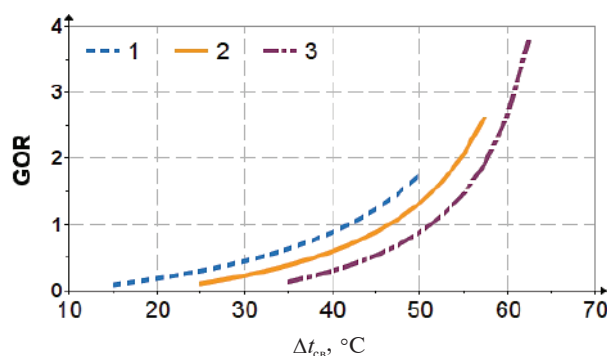


Рис. 3. Залежність GOR від температур солоної води на вході та виході з осушувача: 1 – $t_{1в} = 10$ °C; 2 – $t_{1в} = 20$ °C; 3 – $t_{1в} = 30$ °C

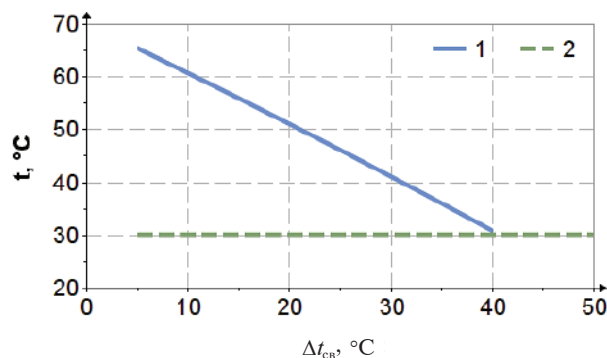


Рис. 4. Залежність температури солоного залишку в зволожувачі від перепаду температур солоної води в осушувачі: 1 – $t_{4в}$; 2 – $t_{1п}$

Як показано на рис. 4, зі збільшенням різниці температур солоної води в осушувачі $\Delta t_{св}$ температура солоного залишку в зволожувачі зменшуватиметься доти, поки не зрівняється з температурою $t_{1п}$. За подальшого збільшення $\Delta t_{св}$ неможлива коректна робота установки.

На практиці необхідно, щоб різниця між $t_{4в}$ та $t_{1п}$ була не більше ніж $5\text{ }^{\circ}\text{C}$. На рис. 3 показано, що за $t_{1в} = 10\text{ }^{\circ}\text{C}$ максимальне значення температури солоної води на виході з установки $t_{2в}$ становить $50\text{ }^{\circ}\text{C}$, а у разі $t_{1в} = 30\text{ }^{\circ}\text{C}$ – $62\text{ }^{\circ}\text{C}$. Очевидно, що догріти солону воду до максимально можливої температури ($70\text{ }^{\circ}\text{C}$) без додаткового нагрівника у цьому циклі нереально.

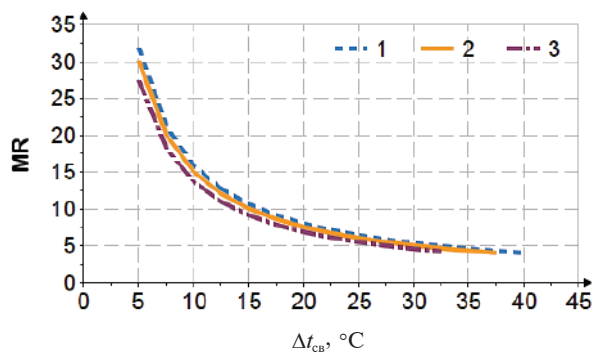


Рис. 5. Залежність MR від перепаду температур солоної води в осушувачі: 1 – $t_{1в} = 10\text{ }^{\circ}\text{C}$; 2 – $t_{1в} = 20\text{ }^{\circ}\text{C}$; 3 – $t_{1в} = 30\text{ }^{\circ}\text{C}$

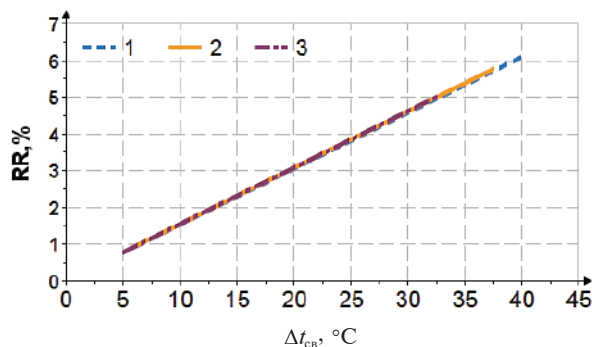


Рис. 6. Залежність RR від перепаду температур солоної води в осушувачі: 1 – $t_{1в} = 10\text{ }^{\circ}\text{C}$; 2 – $t_{1в} = 20\text{ }^{\circ}\text{C}$; 3 – $t_{1в} = 30\text{ }^{\circ}\text{C}$

Як бачимо з рис. 5 та 6, коефіцієнти відновлення води масової витрати фактично не залежать від $t_{1в}$. Коефіцієнт MR зі збільшенням $\Delta t_{св}$ зменшується. Це означає, що чим більша температура $t_{2в}$, тим більше повітря треба подати до осушувача, щоб сконденсувати більше прісної води з повітря. Коефіцієнт відновлення води RR зі збільшенням $\Delta t_{св}$ теж збільшується. Це означає, що чим більша температура $t_{2в}$, тим більше прісної води можна отримати в установці.

Як бачимо з рис. 7, чим більша температура $t_{3в}$, тим ефективність установки вища. Це зумовлено тим, що зі збільшенням температури солоної води на вході до зволожувача $t_{3в}$ темпе-

ратура повітря на виході зі зволожувача $t_{2п}$ теж збільшуватиметься (див. рис. 2, а). Таким чином у зволожувачі до повітря підводиться більше теплоти і вологи (див. формулу (4)). Це дає змогу отримати більше прісної води в осушувачі. Слід зауважити, що на практиці верхнє значення температури $t_{3в}$ слід обмежувати ($50\text{--}70\text{ }^{\circ}\text{C}$), щоб запобігати утворенню накипу.

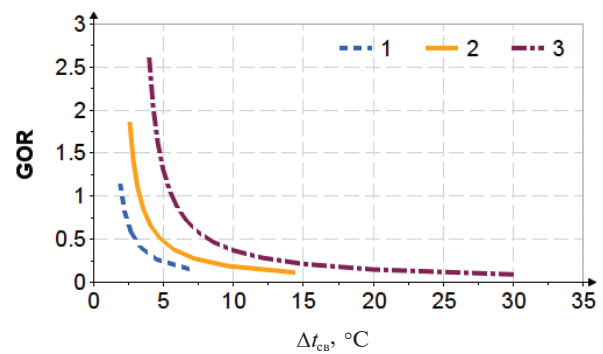


Рис. 7. Залежність GOR від MR: 1 – $t_{3в} = 50\text{ }^{\circ}\text{C}$; 2 – $t_{3в} = 60\text{ }^{\circ}\text{C}$; 3 – $t_{3в} = 70\text{ }^{\circ}\text{C}$

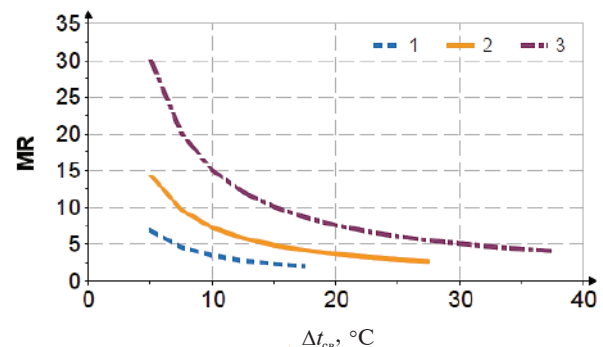


Рис. 8. Залежність RR від перепаду температур солоної води в осушувачі: 1 – $t_{1в} = 10\text{ }^{\circ}\text{C}$; 2 – $t_{1в} = 20\text{ }^{\circ}\text{C}$; 3 – $t_{1в} = 30\text{ }^{\circ}\text{C}$

Ефективність установки (GOR) не залежить від температури та вологості повітря на вході до зволожувача через те, що параметри зовнішнього повітря не впливають на кількість рідини, яка буде поглинена повітрям у зволожувачі. Ефективність осушувача та зволожувача під час моделювання не змінювалася і становила $\varepsilon_{ос} = 0.965$ і $\varepsilon_{зв} = 0.73$ відповідно.

Висновки

Ефективність установки (GOR) не залежить від параметрів повітря на вході до зволожувача. Це дає змогу замкнути повітряний контур

установки (повітря з виходу осушувача одразу подавати до зволожувача, $t_{1п} = t_{2п}$) й уникнути постійного контролю за забрудненням повітря.

Щоб отримати максимальні значення GOR та RR (які для цього циклу становлять GOR = 3,75 та RR = 6%, див. рис. 3 та 6), потрібно, щоб температура $t_{1в}$ була мінімальною (10–40 °C), а перепад температур в осушувачі $\Delta t_{св}$ і температура солоної води на вході до зволожувача $t_{3в}$ – максимально можливими.

Основними чинниками, які лімітують подальше збільшення GOR та RR, є обмеження максимальних значень $t_{3в}$ та $\Delta t_{св}$. Це зумовлює необхідність використання у циклі нагрівника із зовнішнім підведенням теплоти.

Найбільш ефективним способом підвищення GOR є застосування альтернативної енергії (сонце, димові гази тощо) для підігрівання солоної води у нагрівнику. Таке рішення значно зменшить витрати на теплову енергію $Q_{нар}$ і на порядок підвищить ефективність установки (GOR).

Подальші дослідження будуть спрямовані на створення математичних моделей усіх типів опріснювальних систем із циклом зволоження – осушення повітря для порівняння їх термодинамічної та ексергетичної ефективності; виготовлення експериментальної установки для проведення дослідження гідродинаміки, тепло- і масообміну в зволожувачі та осушувачі з для підвищення їх ефективності.

References

- [1] F. Macedonio et al., “Efficient technologies for worldwide clean water supply,” *Chem. Eng. Process.*, vol. 51, pp. 2–17, 2012, doi: 10.1016/j.cep.2011.09.011.
- [2] “Nature-based solutions for water,” UNESCO, Paris, France, 2018. [Online]. Available: <https://unesdoc.unesco.org/ark:/48223/pf0000261424>.
- [3] T. Luo et al., “Aqueduct projected water stress country rankings,” World Resources Institute, Washington, D.C., USA, 2015. [Online]. Available: <https://www.wri.org/data/aqueduct-projected-water-stress-country-rankings>.
- [4] “Форум з екології та сталого розвитку України 2021.” dpss.gov.ua. <https://dpss.gov.ua/news/ukrayina-mozhe-zitknutisya-z-deficitom-yakisnoyi-pitnoyi-vodi-vladislava-magalecka> (accessed Jan. 28, 2021).
- [5] A. Subramani and J. G. Jacangelo, “Emerging desalination technologies for water treatment: a critical review,” *Water Res.*, vol. 75, pp. 164–187, 2015, doi: 10.1016/j.watres.2015.02.032.
- [6] Y. B. Karhe and P. V. Walke, “A solar desalination system using humidification–dehumidification process: a review of recent research,” *Int. J. Mod. Eng. Res.*, vol. 3, no. 2, pp. 962–969, 2013.
- [7] A. S. A. Mohamed et al., “Desalination process using humidification–dehumidification technique: a detailed review,” *Int. J. Energy Res.*, vol. 45, no. 3, pp. 3698–3749, 2020, doi: 10.1002/er.6111.
- [8] “Dynamic growth for desalination and water reuse in 2019.” IDADESAL.org. <https://idadesal.org/dynamic-growth-for-desalination-and-water-reuse-in-2019> (accessed Jan. 28, 2021).
- [9] F. Wunder et al., “Numerical simulation of heat transfer in a horizontal falling film evaporator of multiple-effect distillation,” *Desalination*, vol. 401, pp. 206–229, 2017, doi: 10.1016/j.desal.2016.09.020.
- [10] Y. Zhang et al., “Application of solar energy in water treatment processes: a review,” *Desalination*, vol. 428, pp. 116–145, 2018, doi: 10.1016/j.desal.2017.11.020.
- [11] Velmurugan et al., “Humidification–dehumidification desalination system: an overview,” *Int. J. S. Res. Sci. Tech.*, vol. 4, no. 5, pp. 1163–1177, 2018.
- [12] I. Tlili et al., “Performance enhancement of a humidification–dehumidification desalination system,” *J. Therm. Anal. Calorim.*, vol. 140, pp. 309–319, 2019, doi: 10.1007/s10973-019-08775-8.
- [13] V. Patelet al., “Experimental and theoretical evaluation of bubbler humidifier for humidification–dehumidification water desalination system,” *Heat Mass Transf.*, vol. 55, pp. 3201–3210, 2019, doi: 10.1007/s00231-019-02659-1.
- [14] D. U. Lawal et al., “Experimental investigation of heat pump driven humidification–dehumidification desalination system for water desalination and space conditioning,” *Desalination*, vol. 475, pp. 114199, 2020, doi: 10.1016/j.desal.2019.114199.
- [15] P. P. Kudelia and S. V. Dubovskiy, “Energy and exergy approaches to problem of rational energy use,” *Enerhetyka: ekonomika, tekhnolohii, ekolohiia*, vol. 2, pp. 7–16, 2020.
- [16] “C++ library of properties for 122 components.” CoolProp.org. <http://www.coolprop.org> (accessed Jan. 28, 2021).

Sereda V.V., Solomakha A.S., Prytula N.O., Zabolotny O.A.

THERMODYNAMIC ANALYSIS OF THERMAL DESALINATION SYSTEM WITH HUMIDIFICATION–DEHUMIDIFICATION CYCLE

Problems. Today desalination of seawater provides a worldwide fresh water production of 24.5 million m³ per day. One of the promising desalination methods is thermal humidification-dehumidification process. In spite of significant benefits this technology, it has a key drawback – high thermal energy consumption.

The aim of the study. Determine the optimal thermodynamic parameters of working substances (air and water) in elements of desalination system with humidification–dehumidification cycle to obtain maximum performance with a minimum energy consumption.

Methodology of implementation. Mathematical modeling of the humidification-dehumidification cycle at different input parameters. The thermodynamic model was created using equations of mass and energy balances of the desalination system.

Research results. System Gain output ratio (GOR) increases with increasing salt water temperature at the outlet of the dehumidifier. Maximum values of GOR and recovery ratio (RR) are observed at the maximum temperature difference (maximum heating) of salt water in the dehumidifier. It is shown that values of GOR and RR increase with increasing salt water temperature at the inlet to the humidifier. It is found that GOR value is independent of the temperature and humidity of the air at the inlet to the humidifier.

Conclusions. The maximum values of GOR and RR for the humidification-dehumidification cycle with additional water heating are 3.75 and 6%, respectively. To obtain such results, the temperature of salt water at the inlet to the system (dehumidifier) should be equal to 10 °C, and the temperature of salt water at the inlet to the humidifier – 70 °C. At the same time, temperature and humidity of the air can have any values. The use of renewable energy for heating salt water in the heater will significantly increase system GOR.

Key words: gain output ratio, humidification-dehumidification, mathematical model, thermal desalination, thermodynamic analysis.

Рекомендована Радою
теплоенергетичного факультету
КПІ ім. Ігоря Сікорського

Надійшла до редакції
17 грудня 2021 року

Прийнята до публікації
14 лютого 2022 року

DOI: 10.20535/kpissn.2021.4.252054

УДК 621.574.9

Я.Є. Трокоз*, П.О. Барабаш, О.Б. Голубев, П.П. Куделя

*corresponding author: yar_t@ukr.net

РОЗРОБЛЕННЯ ТА ДОСЛІДЖЕННЯ ХОЛОДИЛЬНОГО ЦИКЛУ З БІАГЕНТНИМ СТРУМИННИМ КОМПРЕСОРОМ

Проблематика. Нині в пароежекторних холодильних машинах (ПЕХМ) використовують моноагентні струминні компресори, робочим тілом яких є переважно вода. Попри низьку ефективність, ПЕХМ мають низьку перевагу порівнянні з холодильними машинами, що використовують механічні компресори: невелику собівартість, простоту виготовлення та надійність роботи. Недоліком ПЕХМ є неможливість досягнення цими установками температур, нижчих за 0 °С. Це обмеження можна подолати, якщо використати біагентний (вода-фреон) струминний компресор (БАСК) і реалізувати новий парокompresійний холодильний цикл.

Мета дослідження. Експериментально підтвердити працездатність холодильної машини, що працює за парокompresійним циклом з використанням біагентного (вода-фреон) струминного компресора. Визначити ефективність роботи експериментальної холодильної машини і знайти оптимальні геометричні характеристики БАСК.

Методика реалізації. Було створено експериментальний стенд для дослідження холодильної установки з використанням БАСК. За розробленою методикою визначався холодильний коефіцієнт ефективності установки ϵ .

Результати дослідження. Отримано експериментальні значення холодильного коефіцієнта машини ϵ залежно від зміни коефіцієнта ежекції q БАСК. Максимальні значення холодильного коефіцієнта установки $\epsilon_{\max} = 1,3 \div 1,4$ відповідають значенням коефіцієнта ежекції $q_{\text{opt}} \approx 0,25$, за яких також досягається максимальний коефіцієнт корисної дії БАСК $\eta_{\max} = 0,35 \div 0,4$

Висновки. Експериментально підтверджено працездатність холодильної машини, що працює за парокompresійним циклом із використанням біагентного (вода-фреон) струминного компресора. Задля розширення температурного діапазону циклу в бік пониження температур кипіння у випарнику ($t_{\text{кип}} \leq -20$ °С), заплановано модернізувати лабораторний стенд застосуванням водних розчинів солей як холодоносія.

Ключові слова: холодильна машина; робоче тіло; біагентний струминний компресор; ежектор; холодильний коефіцієнт; об'ємний коефіцієнт ежекції.

Вступ

Холодильна машина – це пристрій, що працює за зворотним (холодильним) термодинамічним циклом і слугує для охолодження об'єктів до температури, нижчої за температуру довкілля. Охолодження відбувається завдяки тому, що енергія у формі теплоти переноситься робочим тілом (фреони, вода, повітря, гази, водні розчини солей, тверді матеріали тощо) від менш до більш нагрітих об'єктів. Холодильний цикл реалізується за витрати зовнішньої енергії.

Нині найпоширенішими є такі типи холодильних машин: компресійні, паро-ежекторні, абсорбційні і термоелектричні (залежно від застосованого робочого тіла і холодильного циклу) [1].

Установка з БАСК за принципом дії схожа на наявні установки струминного типу – ПЕХМ. У ПЕХМ рідка фаза робочої речовини (якою найчастіше є вода) перетворюється на пару у випарнику, поглинаючи відведену від охолоджуваного об'єкта теплоту. Далі ця пара (пасивний потік) відсмоктується та стискається струминним паровим компресором (ежектором), ак-

Рекомендуємо цитувати цю статтю так: Я.Є. Трокоз, П.О. Барабаш, О.Б. Голубев, П.П. Куделя, “Розроблення та дослідження холодильного циклу з біагентним струминним компресором”, *Наукові вісті КНУ*, № 4, с. 77–82, 2021. doi: 10.20535/kpissn.2021.4.252054.

Please cite this article as: Ya.Ye. Trokoz, P.O. Barabash, O.B. Golubev and P.P. Kudelya, “Refrigeration cycle development and research with biagent jet compressor”, *KPI Science News*, no. 4, pp. 77–82, 2021. doi: 10.20535/kpissn.2021.4.252054.

тивним потоком якого є пара, що генерується в парогенераторі завдяки підводу теплоти від зовнішнього джерела. Парова суміш з ежектора надходить у конденсатор, де охолоджується і конденсується. Конденсат частково повертається у випарник, частково у парогенератор.

Перевагами ПЕХМ є простота конструкції та надійність роботи струминного компресора, а також низька його ціна в порівнянні, наприклад, із механічними компресорами (енергія підводиться у формі роботи) парокompресорних холодильних установок. Ще однією перевагою є використання як робочого тіла води, що широко розповсюджена та має низьку ціну в порівнянні з фреонами.

Недоліком установок цього типу є велика витрата води, яку використовують для охолодження пари, оскільки необхідно охолоджувати не тільки холодоагент, а й робочу пару, яку підводять до струминного компресора. До того ж застосування води як холодоагента для досягнення температур, близьких до 0 °С, необхідно підтримувати надзвичайно низький тиск у випарнику (близько 600 Па), а здійснити охолодження до температур, нижчих за 0 °С, неможливо.

Щоби подолати ці недоліки, у НТУУ “КПІ” було розроблено парокompресійний холодильний цикл, де застосовано біагентний (вода–фреон) струминний компресор [2]. Використання цього циклу дає можливість отримати низьку перевагу у порівнянні з відомими ПЕХМ, а саме:

- отримання у випарнику температури, нижчої за 0 °С;
- необов’язковість парогенератора високих параметрів для роботи струминного компресора та, відповідно, додаткового конденсатора робочої пари, що значно знижує вартість отримання холоду;
- непотрібність надзвичайно низького тиску у випарнику, що суттєво знижує його масогабаритні показники.

Постановка задачі

Експериментально підтвердити працездатність холодильної машини, що працює за парокompресійним циклом із використанням біагентного (вода–фреон) струминного компресора. Визначити ефективність роботи експериментальної холодильної машини та знайти оптимальні геометричні характеристики струминного компресора.

Опис біагентного струминного компресора

Головною ідеєю винаходу [2] є заміна робочого тіла (води) в пароежекторному циклі на низькокиплячі органічні робочі тіла (фреони), а також використання води як активного потоку в струминному компресорі для забезпечення роботи циклу. Як холодильний агент вибрано фреон R-141b, що практично не розчиняється у воді, а за температури 20 ÷ 30 °С має близький до атмосферного тиск насичення та, відповідно, є безпечним для зберігання; за впливом на озоновий шар належить до холодоагентів групи В (спричиняє слабке виснаження озонового шару), а за дією на організм людини є шкідливою речовиною 4-го класу (малонебезпечною речовиною).

Основним елементом запропонованого холодильного циклу є біагентний (вода–фреон) струминний компресор. З урахуванням можливостей лабораторної бази для розрахунку прийнято такі вихідні умови:

- тиск води на вході в струминний компресор – $p_1 = 220$ кПа;
- тиск на виході зі струминного компресора – $p_2 = 100$ кПа;
- об’ємна витрата води на вході в струминний компресор – $\dot{V}_в = 2,8 \cdot 10^{-4}$ м³ / с;
- тиск ежектованого середовища (R-141b) – $p_e = 25$ кПа.

Розрахунок геометричних характеристик дослідного БАСК виконано за методикою, описаною в [3]. За отриманими геометричними характеристиками біагентного (вода–фреон) струминного компресора створено його віртуальну модель у програмному забезпеченні “SolidWorks” [4].

На рис. 1 зображено отриману віртуальну модель БАСК, основними елементами якого є:

- приймальна камера;
- звужувальне сопло з вихідним діаметром $d = 4,4 \cdot 10^{-3}$ м;
- циліндрична камера змішування завдовжки $90 \cdot 10^{-3}$ м із внутрішнім діаметром $d = 8 \cdot 10^{-3}$ м;
- дифузор завдовжки $70 \cdot 10^{-3}$ м із кутом розкриття $\alpha = 8^\circ$.

Модуль пристрою становить

$$A = \left(\frac{f_{3M}}{f_c} \right) = 3,3,$$

де f_{3M} – площа поперечного перерізу камери змішування, м²; f_c – площа поперечного перерізу сопла, м².

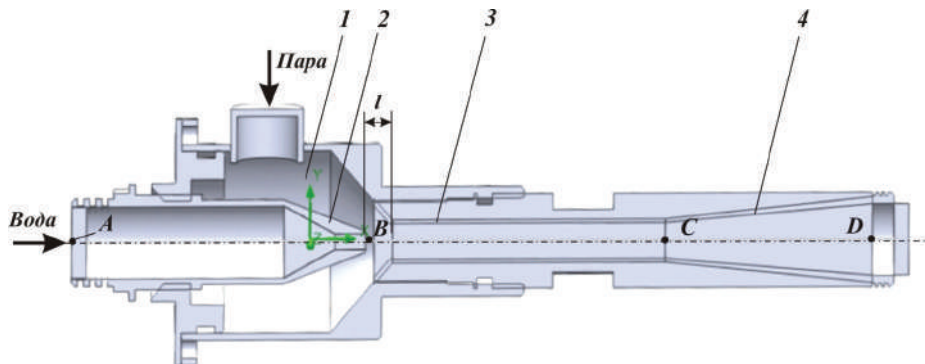


Рис. 1. Біагентний струминний компресор: 1 – приймальна камера; 2 – звузувальне сопло; 3 – циліндрична камера змішування; 4 – дифузор

Задля візуалізації процесів у пристрої фізичну модель БАСК виготовлено з прозорого матеріалу (полікарбонату).

8, де віддає тепло киплячому фреону і надалі, охолоджений, поступає у збірник 12.

Експериментальний стенд

На рис. 2 зображено принципову схему експериментальної установки для дослідження парокompресійного холодильного циклу з використанням біагентного (вода–фреон) струминного компресора.

Принцип роботи експериментальної холодильної установки такий. Циркуляційна помпа 2 подає воду до водопарового струминного компресора 1, який відсмоктує пару фреону з випарника 8. Водофреонова суміш надходить у сепаратор-охолоджувач 3, де відбувається розділення рідкого фреону та води. Потім вода через ротаметр 4 і циркуляційну помпу 2 знову подається до струминного компресора 1, а рідкий фреон, що накопичується в нижній частині сепаратора 3, потрапляє через ротаметр 6 до дроселя 9, де відбувається дроселювання з пониженням тиску та температури. Після дроселя фреон потрапляє до випарника 8, де відбувається його кипіння з відбором теплоти від холодоносія – підведення теплоти в цикл (Q_r). Відведення теплоти в цикл (Q_x) відбувається у сепараторі-охолоджувачі 3. Холодоносій зі збірника 12 подається помпою 11 до електричного підігрівача 10, звідки холодоносій через ротаметр 7 поступає у випарник

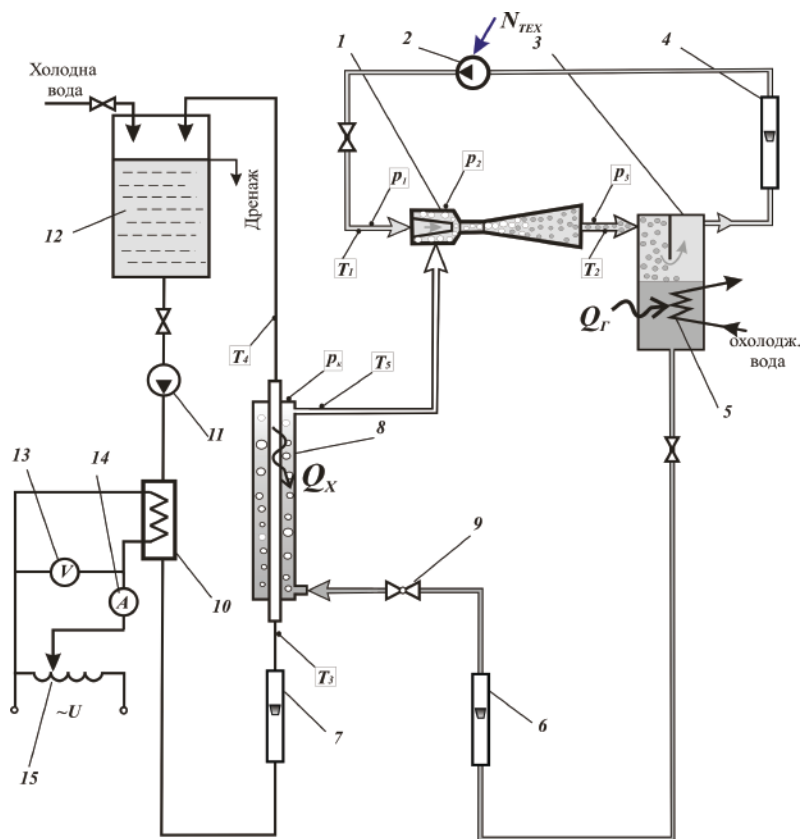


Рис. 2. Схема холодної установки: 1 – біагентний струминний компресор; 2 – помпа контуру активного потоку струминного компресора; 3 – сепаратор-охолоджувач; 4, 6, 7 – ротаметри; 5 – охолоджувальний контур (відвід теплоти з циклу); 8 – випарник, 9 – дросель; 10 – електричний підігрівач холодоносія; 11 – циркуляційна помпа контуру холодоносія; 12 – збірник холодоносія; 13 – вольтметр; 14 – амперметр; 15 – автотрансформатор; T_1, T_2, T_3, T_4, T_5 – точки вимірювання температури; p_1, p_2, p_3, p_4, p_5 – точки вимірювання статичного тиску

Методика розрахунку основних величин

Ефективність запропонованого холодильного циклу оцінювалася за допомогою визначення експериментального холодильного коефіцієнта:

$$\varepsilon = \frac{Q_X}{N_{\text{ц}}},$$

де Q_X – теплове навантаження випарника, Вт; $N_{\text{ц}}$ – витрачена потужність у циклі, Вт.

Тепловий потік у випарнику визначався двома способами (різниця значень не перевищувала $\pm 5\%$):

а) за зміною ентальпії холодоносія

$$Q_X = G_b \cdot c_b (t'_b - t''_b),$$

де G_b – витрата холодоносія (води) у випарнику, кг/с; c_b – питома теплоємність води, Дж/кгК; t'_b, t''_b – температури холодоносія на вході та виході з випарника відповідно;

б) за витратою холодоагента у випарнику

$$Q_X = G_{\phi} \cdot r_{\phi},$$

де $G_{\phi} = \dot{V}_{\phi} \cdot \rho_{\phi}$ – масова витрата холодоагента, кг/с; \dot{V}_{ϕ} – об'ємна витрата рідкої фази холодоагента, м³/с; ρ_{ϕ} – густина рідкої фази холодоагента, кг/м³; r_{ϕ} – теплота пароутворення фреону, кДж/кг.

Потужність, витрачену на здійснення циклу, визначали за рівнянням

$$N_{\text{ц}} = \frac{\dot{V}_b \cdot (p_1 - p_2)}{\eta_{\text{ц}}},$$

де \dot{V}_b – витрата води (активний потік) у струминному компресорі, м³/с; p_1 і p_2 – тиск активного потоку на вході та виході струминного компресора відповідно, Па; $\eta_{\text{ц}}$ – ККД циркуляційної помпи, що забезпечує перепад тиску ($p_1 - p_2$).

Похибка визначення експериментального холодильного коефіцієнта ε_e не перевищує $\pm 15\%$.

В експериментах для кожного режиму також визначено об'ємний коефіцієнт ежекції $q_v = \dot{V}_{\phi} / \dot{V}_b$, масовий коефіцієнт ежекції $q_g = G_{\phi} / G_b$, коефіцієнт напору $h = (p_2 - p_e) / (p_1 - p_e)$ та ККД струминного компресора $\eta = q_v h / (1 - h)$.

Результати експериментального дослідження

Попередні тестування струминного компресора в автономному режимі (за відсутності ежектованої пари) дали змогу визначити відстань

від сопла до камери змішування l (див. рис. 1), на якій можна досягти найбільшого розрідження у приймальній камері. Результати тестування наведено на рис. 3.

Найбільше розрідження досягнуто на відстані від зрізу сопла до камери змішування $l \approx 9 \cdot 10^{-3}$ м. Цей результат відповідає діапазону, рекомендованому в літературі для струминних компресорів [3].

Віртуальні дослідження моделі струминного компресора в програмному середовищі “SolidWorks Simulation” дали можливість визначити розподіл тиску активного потоку (води) по довжині проточної частини струминного компресора. Результати віртуального та фізичного експерименту з реальним дослідним струминним компресором наведено на рис. 4. Дані щодо зміни тиску дали можливість визначити характер протікання процесів із холодоагентом у проточній частині струминного компресора та побудувати цикл, досліджений у $T - s$ і $\ln p - h$ координатах (див. рис. 5).

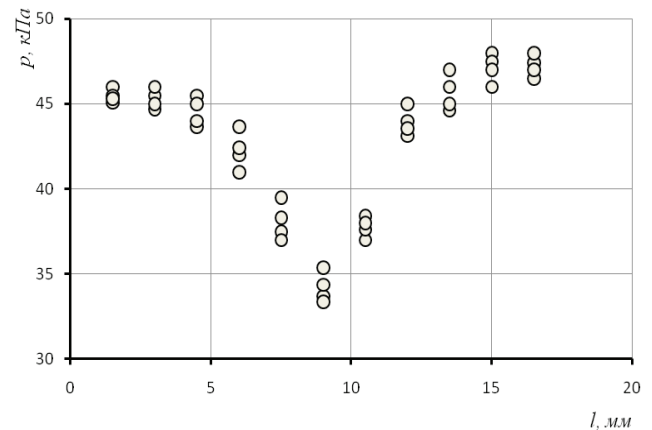


Рис. 3. Залежність абсолютного тиску середовища в приймальній камері біагентного струминного компресора від відстані між зрізом сопла та камерою змішування ($\dot{V}_b = 2,8 \cdot 10^{-4}$ м³/с)

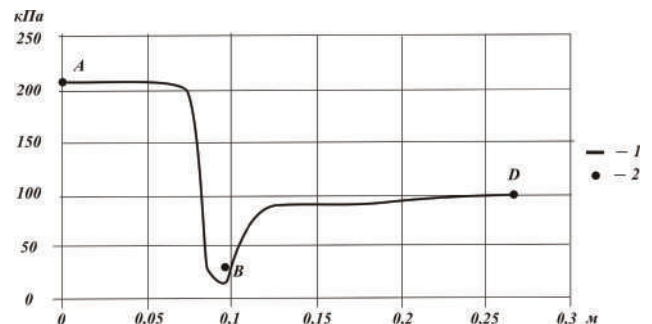


Рис. 4. Розподіл тиску середовища по довжині струминного компресора: 1 – розрахунок віртуальної моделі в програмному середовищі “SolidWorks Simulation”; 2 – експеримент

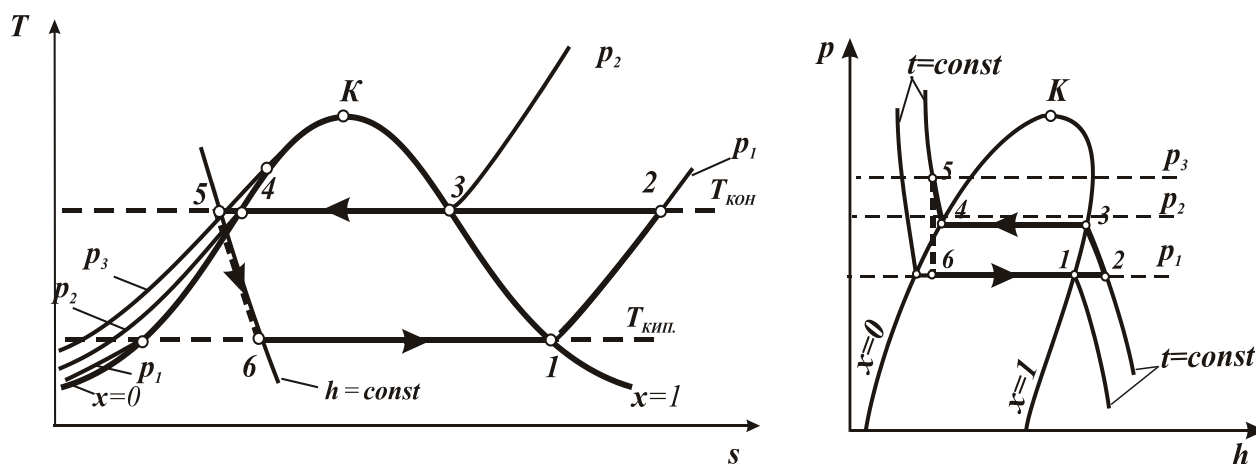


Рис. 5. Термодинамічний цикл холодильної установки з біагентним струминним компресором у $T - s$ і $\ln p - h$ координатах

На основі візуальних спостережень за рухом водофреонового потоку в струминному компресорі встановлено, що парова фаза фреону відразу після зрізу сопла (точка B на рис. 1) утворює зі струменем води білу дрібнодисперсну піну. Відбувається безпосередній теплообмін між бульбашками фреону та супутнім струменем води. Наближений розрахунок показав, що в усьому діапазоні експериментальних значень об'ємних коефіцієнтів ежекції $q_v = 0,0018 \div 0,38$ вирівнювання температури мікробульбашки з температурою потоку води відбувається на шляху, меншому ніж $0,001$ м. За браком швидкісної камери не вдалося виявити початок зони схлопування (конденсації) бульбашок пари фреону при підвищенні тиску водофреонового потоку в дифузори струминного компресора.

Ураховуючи малі розміри бульбашок і малі значення масових коефіцієнтів ежекції $q_g = 6,1 \cdot 10^{-6} \div 8,6 \cdot 10^{-4}$, можна вважати, що тиск фреону синхронно змінюється з тиском водяного потоку, температура якого практично не змінюється. Отже процеси, що відбуваються у запропонованому циклі, можна описати так.

Пара фреону із випарника подається у струминний компресор (див. точка B на рис. 1 та точка 1 на рис. 5). Практично миттєво відбувається ізобарний нагрів пари фреону до температури потоку води (див. точка 2 на рис. 5). Далі пара фреону ізотермічно стискається до тиску насичення, і починається конденсація фреону (див. точка 3

на рис. 5). На рис. 1, початок конденсації знаходиться на проміжку між точками B та C . На проміжку між точками C та D (див. рис. 1) закінчується конденсація (у межах дифузора струминного компресора). Далі відбувається стиснення вже конденсату фреону до значення тиску на виході з дифузора (див. точка D на рис. 1 та точка 5 на рис. 5). На виході з дифузора утворюється гетерогенна суміш рідкого фреону і води. У сепараторі рідкий фреон відділяється від води і через дросель (процес 5–6 рис. 5) подається у випарник, де кипить (процес 6–1 рис. 5).

На рис. 6 показано залежність холодильного коефіцієнта експериментальної установки ϵ_e зі струминним компресором від об'ємного коефіцієнта ежекції q_v :

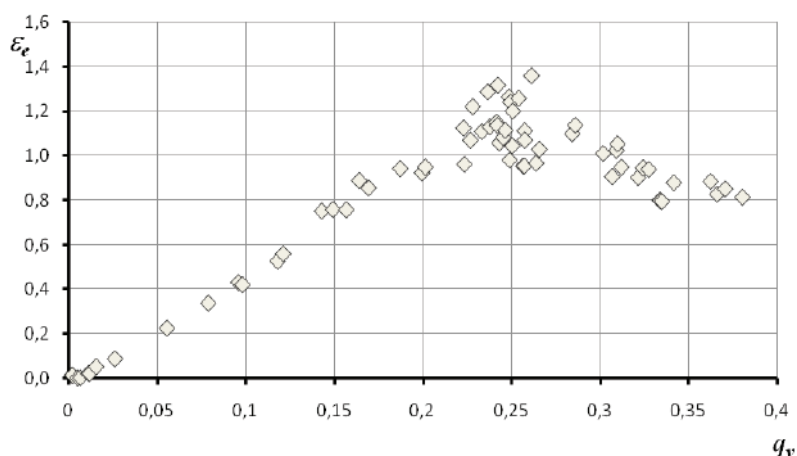


Рис. 6. Залежність холодильного коефіцієнта експериментальної установки ϵ_e зі струминним компресором від об'ємного коефіцієнта ежекції q_v .

Обговорення результатів

Максимальні значення експериментально-го холодильного коефіцієнта установки $\varepsilon_{\max} = 1,3 \div 1,4$ відповідають значенням коефіцієнта ежекції $q_{\text{онм}} \approx 0,25$, за яких також досягається максимальний ККД біагентного (вода–фреон) струминного компресора $\eta_{\max} = 0,35 \div 0,4$. Більшість експериментів проведено в діапазоні температур: температура конденсації $t_{\text{кон}} = 20$ °С, температура кипіння $t_{\text{кип}} = -2$ °С. Застосування нижчих температур кипіння у випарнику було обмежене браком можливості використати незамерзний холодоносіє, що могло б призвести до розморожування випарника. Але під час налагоджувальних експериментів мінімальні температури кипіння у випарнику короткочасно досягали (до моменту замерзання холодоносія) $t_{\text{кип}} = -20$ °С.

References

- [1] T.V. Moroziuk, *Theory of refrigerators and heat pumps*. Odessa, Ukraine: Studio “Negotsiant”, 2006.
- [2] Ya.Ye. Trokoz, P.O. Barabash, and O.B. Golubev, “The method of obtaining cold in the binary refrigeration cycle “KPI”, UA Patent. 72697, Aug. 27, 2012.
- [3] E. Ya. Sokolov, and N.M. Zinger, *Inkjet devices*. Moscow, USSR: Energoatomizdat, 1989.
- [4] A.A. Alyamovskiy, *SolidWorksSimulation. How to solve practical problems*. St. Petersburg, Russia: BHV-Petersburg, 2012.

Ya.Ye. Trokoz, P.O. Barabash, O.B. Golubev, P.P.Kudelya

DEVELOPMENT AND RESEARCH OF REFRIGERATION CYCLE WITH BIAGENT JET COMPRESSOR

Background. Currently, in the cycles of steam-ejector refrigeration machines (SERM) are used monoagent jet compressors, the working body in which is water. Despite their low efficiency, SERM have a number of advantages over refrigeration machines that use mechanical compressors: low cost, ease of manufacture and reliability. The use of SERM is largely limited by the inability of these installations to reach temperatures below 0 °С. This limitation can be overcome by applying a fundamentally new steam compression cycle using a bioagent (water-freon) jet compressor (BAJC).

Objective. The purpose of the paper is experimentally confirm the operability of a refrigeration machine operating on a steam compression cycle using a bioagent (water-freon) jet compressor. Determine the efficiency of the experimental refrigeration machine and find the optimal geometric characteristics of the jet compressor.

Methods. An experimental stand was created to study the refrigeration plant using BAJC. The refrigeration machine efficiency ε was determined according to the developed method.

Results. Experimental values of the cooling machine coefficient ε were obtained depending on the change in the ejection coefficient q of BAJC. The maximum machine coefficient values $\varepsilon_{\max} = 1,3 \div 1,4$ correspond to those of the ejection coefficient $q_{\text{opt}} \approx 0,25$, which also achieves the maximum efficiency $\eta_{\max} \approx 0,35 \div 0,45$ of the BAJC.

Conclusions. The efficiency of a refrigeration unit operating on a steam compression cycle using a bioagent (water-freon) jet compressor has been experimentally confirmed. To expand the temperature spectrum of the cycle in the direction of lowering the boiling points in the evaporator ($t_{\text{boil}} \leq -20$ °С), it is planned to modernise the laboratory bench by using aqueous salt solutions as a coolant.

Key words: refrigeration unit, working fluid, biagent jet compressor, ejector, refrigeration coefficient, ejection volume coefficient.

Рекомендована Радою
теплоенергетичного факультету
КПІ ім. Ігоря Сікорського

Надійшла до редакції
08 грудня 2021 року

Прийнята до публікації
14 квітня 2022 року

DOI: 10.20535/kpissn.2021.4.245103

УДК621.833.16

І. В. Янчевський*, М. В. Нехороших

КПІ ім. Ігоря Сікорського, Київ, Україна

*correspondingauthor:i.yanchevskiy@kpi.ua

ВПЛИВ ГЕОМЕТРИЧНИХ ПАРАМЕТРІВ КРИВОШИПНО-ПЛАНЕТАРНОГО РЕДУКТОРА НА ЦИКЛОГРАМУ РОБОТИ КОНТАКТНИХ ПАР

Проблематика. Кривошипно-планетарні редуктори (КПР) мають низку переваг: компактність, високу навантажувальну здатність, високу жорсткість конструкції та ін. Водночас немає регламентованої методології їх кінематичного та силового розрахунку, що обмежує їх розроблення та серійне виробництво. Створення ефективних методів розрахунку цих редукторів задля покращення їх експлуатаційних показників на етапі проектування має важливе прикладне значення.

Мета дослідження. Розробити аналітичний метод визначення кута тиску та значення контактної сили в контактних парах “отвір сателіта – цівка вихідної ланки” КПР залежно від кута між кривошипами його колінчастого вала (V-подібної форми), а також від кількості сателітів і цівок вихідної ланки. Застосовуючи розроблений метод, перевірити можливості проектування КПР з довільними значеннями зазначених параметрів і дослідити їх вплив на циклограму роботи контактних пар і максимальне значення контактної сили з урахуванням нелінійності деформування матеріалів.

Методика реалізації. На першому етапі на основі формул аналітичної геометрії показано можливість реалізації КПР із довільним значенням кута між кривошипами колінчастого вала, а також довільною кількістю сателітів і цівок вихідної ланки. Проаналізовано значення кутів тиску в контактних парах “отвір сателіта – цівка вихідної ланки” та побудовано циклограму роботи ведучих контактних пар. На другому етапі з використанням принципу можливих переміщень отримано формулу для розрахунку контактної сили у ведучих контактних парах як функції кута повороту вхідного вала. Під час обчислення контактної сили використано нелінійну модель деформування матеріалу.

Результати дослідження. Розроблено аналітичний метод побудови циклограми роботи контактних пар КПР і розрахунку контактних сил залежно від геометричних параметрів редуктора. Показано можливість реалізації редукторів із довільним значенням кута між кривошипами колінчастого вала, а також довільною кількістю сателітів і цівок вихідної ланки.

Висновки. Отримано значення контактної сили – функцію кута повороту колінчастого вала, кількості сателітів і цівок вихідної ланки. За незмінної кількості одночасно активних ведучих контактних пар і, відповідно, рівномірної передачі навантаження між ними закон зміни контактної сили за цикл подібний до параболічного. В іншому випадку зміна контактної сили за цикл має досить складний характер. Отримані результати можуть бути використані за уточненого розрахунку ККД КПР й ухвалення конструкторських рішень на етапі його проектування.

Ключові слова: кривошипно-планетарний редуктор; геометричні параметри редуктора; циклограма; контактна сила; кут тиску.

Вступ

Механічні редуктори різних типів і розмірів широко застосовують у багатьох машинах. Планетарні редуктори як один із типів механічних

редукторів вирізняються можливістю реалізувати широкий діапазон передатних чисел і мають високий ККД, менші масогабаритні характеристики в порівнянні з рядними передачами [1]. Найбільшого поширення планетарні передачі

Рекомендуємо цитувати цю статтю так: І.В. Янчевський, М.В. Нехороших, “Вплив геометричних параметрів кривошипно-планетарного редуктора на циклограму роботи контактних пар”, *Наукові вісті КПІ*, № 4, с. 83–91, 2021. doi: 10.20535/kpissn.2021.4.245103.

Please cite this article as: I.V. Yanchevskiy, M.V. Niekhoroshykh, “Influence of geometric parameters of crankshaft-planetary reducer on cyclograph of work of contact pairs”, *KPI Science News*, no. 4, pp. 83–91, 2021. doi: 10.20535/kpissn.2021.4.245103.

набули у важконавантажених і відповідальних механічних приводах, у робототехніці та мехатронних системах, де вони практично не мають альтернативи [2], [3].

Кривошипно-планетарні редуктори – один із варіантів конструктивного виконання планетарних редукторів. В основу КПР покладено планетарний механізм типу К-Н-V [4]–[6] із внутрішнім зачепленням з малою різницею чисел зубів сателітів і сонячного колеса. Для більшості конструкцій КПР відмітними особливостями є компактність, висока навантажувальна здатність через брак зовнішнього зачеплення, висока жорсткість конструкції, малий момент інерції валів, низький рівень шуму та вібрації [3]. До основних недоліків КПР, які обмежують обсяги їх серійного виробництва, слід віднести відсутність регламентованої методології їх кінематичного та силового розрахунку, підвищені вимоги до точності виготовлення та монтажу, а також пов'язану з цим остаточну ціну готового виробу. Тому вдосконалення таких редукторів, орієнтоване на зниження вартості їх виробництва, та створення ефективних методів їх розрахунку задля покращення їх експлуатаційних показників на етапі проектування мають важливе прикладне значення.

Класичне виконання КПР [1]–[6] передбачає наявність двох сателітів, які встановлюють на опозитно розташовані проміжні шийки (кривошипи) вхідного колінчастого вала. Водночас таке конструктивне рішення не є єдино можливим. Питання оптимізації геометричних параметрів КПР і їх впливу на експлуатаційні характеристики є відкритим.

Постановка задачі

Розробити аналітичний метод визначення кута тиску та значення контактної сили в контактних парах “отвір сателіта – цівка вихідної ланки” КПР залежно від кута між кривошипами його колінчастого вала (з V-подібною формою), а також кількості сателітів і цівок вихідної ланки. Застосовуючи розроблений метод, перевірити можливості проектування редуктора з довільними значеннями зазначених параметрів і дослідити їх вплив на циклограму роботи контактних пар і максимальне значення контактної сили з урахуванням нелінійності деформування матеріалів.

Розрахункова схема редуктора

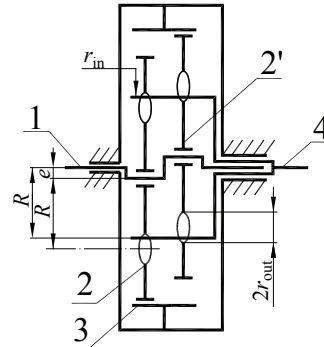


Рис. 1. Структурна схема редуктора: 1 – вхідна ланка; 2, 2' – сателіти; 3 – корпус; 4 – вихідна ланка; e – ексцентриситет; r_{out} – радіус отворів; R – радіус; r_{in} – радіус цівки

Розглянемо КПР, структурну схему якого показано на рис. 1. Ведучою ланкою редуктора є ланка 1, виготовлена як колінчастий вал із кількома кривошипами з ексцентриситетом e . Кількість кривошипів K дорівнює кількості встановлених на них сателітів (позиції 2 та 2'). Кут між кривошипами – α . Загалом значення K й α можуть відрізнятись від класичних ($K = 2$; $\alpha = 180^\circ$). Зубчасті вінця сателітів і корпусу 3 утворюють циліндричні передачі внутрішнього зачеплення. У тілі кожного сателіта є N отворів із радіусом r_{out} , рівномірно розподілених по колу радіусом R . За допомогою поверхонь цих отворів приводиться в обертання вихідна ланка 4 через цівки радіусом r_{in} , які входять в отвори сателітів. Зазначимо, що кількість отворів N у кожному сателіті зазвичай покладають рівною 3 або 4.

Методика розв'язання

Уведемо в розгляд декартові системи координат $A_1x_1y_1$ й $A_1\xi_1\eta_1$, початок яких збігається з центром першого сателіта 2. Уважаємо, що система координат $A_1x_1y_1$ рухається поступально відносно інерціальної системи OXY , вісь Z якої є віссю обертання вхідного (колінчастого) вала 1 і вихідної ланки 4. Інша система координат пов'язана з сателітом, і, відповідно, кут між осями A_1y_1 й $A_1\eta_1$ становитиме φ/u , де φ – кут повороту вхідного вала; u – передатне число редуктора (рис. 2). У початковий момент часу ($\varphi = 0$) осі ординат зазначених вище систем координат збігаються.

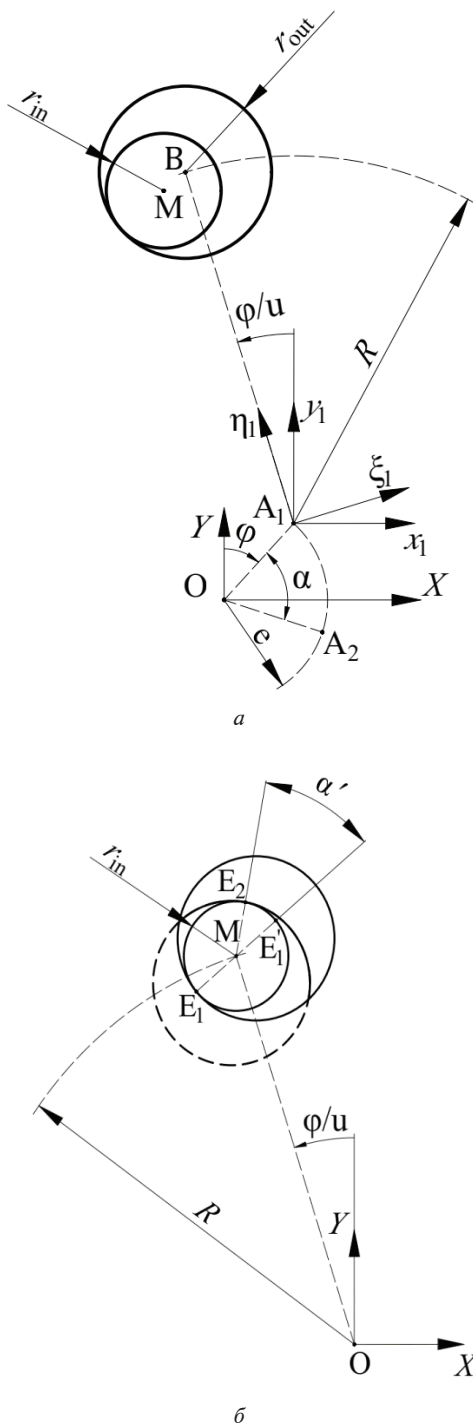


Рис. 2. Системи координат

Положення першого сателіта 2 задамо так, що координата центра одного з його отворів у системі координат $A_1\xi_1\eta_1$ дорівнює $(\xi_{1B}; \eta_{1B}) = (0, R)$. Тоді в системі $A_1x_1y_1$ координати цього ж центру визначають за допомогою матриці повороту систем координат [8]:

$$\begin{bmatrix} x_{1B} \\ y_{1B} \end{bmatrix} = \begin{bmatrix} \cos(\varphi/u) & -\sin(\varphi/u) \\ \sin(\varphi/u) & \cos(\varphi/u) \end{bmatrix} \cdot \begin{bmatrix} \xi_{1B} \\ \eta_{1B} \end{bmatrix}.$$

Так само можна знайти координати центра цього отвору в нерухомій системі координат OXY :

$$\begin{aligned} \begin{bmatrix} X_B \\ Y_B \end{bmatrix} &= \begin{bmatrix} e \cdot \sin(\varphi) \\ e \cdot \cos(\varphi) \end{bmatrix} + \begin{bmatrix} x_{1B} \\ y_{1B} \end{bmatrix} = \\ &= \begin{bmatrix} e \cdot \sin(\varphi) - R \cdot \sin(\varphi/u) \\ e \cdot \cos(\varphi) + R \cdot \cos(\varphi/u) \end{bmatrix}, \end{aligned}$$

де e – ексцентриситет вхідного вала ($e = OA_1$).

З урахуванням того, що передатне число від вхідного вала до сателітів дорівнює u , координати точки E_1 (точки дотику цівки вихідної ланки та відповідного отвору першого сателіта (рис. 2, б)) отримуємо з такої системи алгебраїчних рівнянь:

$$\begin{cases} (X_{E_1} - X_B)^2 + (Y_{E_1} - Y_B)^2 = r_{out}^2; \\ (X_{E_1} + R \cdot \sin(\varphi/u))^2 + (Y_{E_1} - R \cdot \cos(\varphi/u))^2 = r_{in}^2, \end{cases}$$

де r_{in} , r_{out} – радіуси цівок вихідної ланки редуктора та отвору сателіта (див. рис. 2).

За допомогою координат точок E_1 і M визначимо координати точки E_2 – точки дотику цієї ж цівки до другого сателіта в нерухомій системі координат:

$$\begin{bmatrix} X_{E_2} \\ Y_{E_2} \end{bmatrix} = \begin{bmatrix} X_M \\ Y_M \end{bmatrix} + \mathbf{L} \cdot \begin{bmatrix} X_M - X_{E_1} \\ Y_M - Y_{E_1} \end{bmatrix}, \quad (1)$$

де $\begin{bmatrix} X_M \\ Y_M \end{bmatrix} = \begin{bmatrix} -R \sin(\varphi/u) \\ R \cos(\varphi/u) \end{bmatrix}$ – координати точки M

(див. рис. 2, б); \mathbf{L} – лінійний оператор повороту $\overline{ME_1}' = \overline{E_1M}$ на кут $\alpha' = \angle E_2ME_1'$ (див. рис. 2, б) проти годинникової стрілки:

$$\mathbf{L} = \begin{bmatrix} \cos(\alpha') & -\sin(\alpha') \\ \sin(\alpha') & \cos(\alpha') \end{bmatrix}.$$

Загалом точка E_2 не буде діаметрально протилежною до E_1 відносно точки M ($-\pi < \alpha' < \pi$).

Знаючи координати точки E_2 як функцію кута повороту вхідного вала φ , на цьому етапі можна параметризувати рівняння руху отвору другого сателіта задля визначення форми його профілю. Для цього необхідно перейти від системи координат OXY до системи $A_2\xi_2\eta_2$, жорстко пов'язаної з центром іншого сателіта (рис. 3).

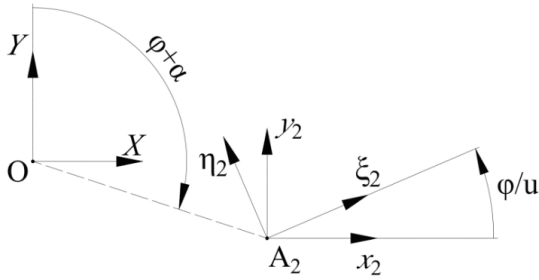


Рис. 3. Перехід до системи координат у центрі другого сателіта

Цей перехід можливий через використання системи координат $A_2x_2y_2$, визначення координат точки E_2 в якій відбувається за законом [8]:

$$\begin{bmatrix} x_{2E_2} \\ y_{2E_2} \end{bmatrix} = \begin{bmatrix} X_{E_2} \\ Y_{E_2} \end{bmatrix} + \begin{bmatrix} \cos(\varphi + \alpha) & \sin(\varphi + \alpha) \\ -\sin(\varphi + \alpha) & \cos(\varphi + \alpha) \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -e \end{bmatrix}.$$

Слід зазначити, що система координат $A_2x_2y_2$ (див. рис. 3), як і $A_1x_1y_1$, рухається поступально відносно інерціальної OXY , однак їх центри збігаються з центрами різних сателітів.

Отже, маючи координати (x_{2E_2}, y_{2E_2}) можна записати параметричне рівняння руху отвору другого сателіта в рухомій системі координат $A_2\xi_2\eta_2$. Ці координати за допомогою класичних тригонометричних співвідношень і матриці повороту

$$\begin{bmatrix} \xi_2 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} \cos(\varphi/u) & \sin(\varphi/u) \\ -\sin(\varphi/u) & \cos(\varphi/u) \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

можна подати як

$$\begin{cases} \xi_{2E_2} = -e \cdot \sin(C_3\varphi + \alpha) + \\ + C_1 \cos(C_3\varphi - \alpha') + C_2 \sin(C_3\varphi - \alpha'); \\ \eta_{2E_2} = R - e \cdot \cos(C_3\varphi + \alpha) + \\ + C_2 \cos(C_3\varphi - \alpha') - C_1 \sin(C_3\varphi - \alpha'); \end{cases} \quad (2)$$

де $C_1 = \frac{1}{2e} \sqrt{2e^2 r_{out}^2 + 2e^2 r_{in}^2 + 2r_{out}^2 r_{in}^2 - e^4 - r_{in}^4 - r_{out}^4}$;

$$C_2 = \frac{r_{out}^2 - r_{in}^2 - e^2}{2e}; \quad C_3 = 1 + \frac{1}{u}.$$

Якщо використати рівність $r_{out} = r_{in} + e$, то система рівнянь (2) суттєво спрощується:

$$C_1 = 0; \quad C_2 = r_{in}; \quad C_3 = 1 + 1/u;$$

$$\xi_{2E_2} = -e \cdot \sin(C_3 \cdot \varphi + \alpha) + r_{in} \cdot \sin(C_3 \cdot \varphi - \alpha'); \quad (3)$$

$$\eta_{2E_2} = R - e \cdot \cos(C_3 \cdot \varphi + \alpha) + r_{in} \cdot \cos(C_3 \cdot \varphi - \alpha').$$

Доведемо, що координати $(\xi_{2E_2}, \eta_{2E_2})$ задовольняють рівняння кола з радіусом r_{out2} з центром у точці $(0, R)$ у рухомій системі координат $A_2\xi_2\eta_2$:

$$\xi_{2E_2}^2 + (\eta_{2E_2} - R)^2 = r_{out2}^2. \quad (4)$$

Підставивши в (4) координати (3) та спростивши вираз, отримуємо

$$r_{out2}^2 = r_{in}^2 + e^2 - 2e \cdot r_{in} \cdot \cos(\alpha + \alpha'). \quad (5)$$

Як видно з рівняння (5), обидві його частини не залежать від параметра φ , що задовольняє фізичну модель. Проте не всі пари (α, α') будуть прийнятними, оскільки отримане коло має бути дотичним до контуру цівки (не має перетинати цей контур). Оскільки ці кути не є функціями кута повороту вхідного вала φ , то для цієї задачі досить, щоб умова (5) виконувалась за $\varphi = 0$. Розглянемо для цього систему координат Mxy у центрі цівки (рис. 4).

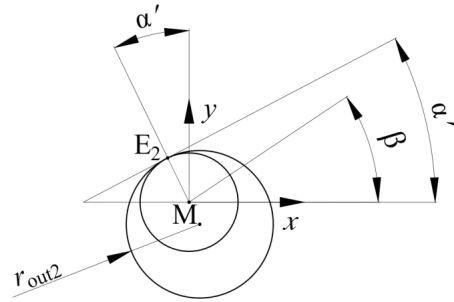


Рис. 4. Контакт поверхні цівки з першим отвором другого сателіта. Параметризуємо рівняння кола контуру цівки за кутом β та знайдемо похідну в точці E_2 ($\beta = \alpha' + \pi/2$):

$$\begin{aligned} \begin{cases} x = r_{in} \cos(\beta) \\ y = r_{in} \sin(\beta) \end{cases} &\Rightarrow \begin{cases} dx/d\beta = -r_{in} \sin(\beta) \\ dy/d\beta = r_{in} \cos(\beta) \end{cases} \Rightarrow \left. \frac{dy}{dx} \right|_{\beta=\alpha'+\pi/2} = \\ &= -\frac{\cos(\alpha' + \pi/2)}{\sin(\alpha' + \pi/2)} = \tan(\alpha'). \end{aligned}$$

Аналогічно запишемо похідну для отвору другого сателіта:

$$\begin{aligned} \begin{cases} \frac{d\xi_{2E_2}}{d\varphi} = C_3 (r_{in} \cos(C_3\varphi - \alpha') - e \cdot \cos(C_3\varphi + \alpha)); \\ \frac{d\eta_{2E_2}}{d\varphi} = C_3 (-r_{in} \sin(C_3\varphi - \alpha') + e \cdot \sin(C_3\varphi + \alpha)). \end{cases} &\Rightarrow \\ \Rightarrow \left. \frac{d\eta_{2E_2}}{d\xi_{2E_2}} \right|_{\varphi=0} &= \frac{r_{in} \cdot \sin(\alpha') + e \cdot \sin(\alpha)}{r_{in} \cdot \cos(\alpha') - e \cdot \cos(\alpha)}. \end{aligned}$$

Умова дотику двох кіл – збіжність кутових коефіцієнтів дотичних:

$$\left. \frac{dy}{dx} \right|_{\beta=\alpha'+\frac{\pi}{2}} = \left. \frac{d\eta_{2E_2}}{d\xi_{2E_2}} \right|_{\varphi=0} \Rightarrow \Rightarrow \frac{r_{in} \cdot \sin(\alpha') + e \cdot \sin(\alpha)}{r_{in} \cdot \cos(\alpha') - e \cdot \cos(\alpha)} = \tan(\alpha').$$

Відповідно до цього рівняння для пари (α, α') отримуємо два можливих варіанти розв'язку:

$$\alpha + \alpha' = 0 \text{ або } \alpha + \alpha' = \pi,$$

які відповідають радіусам

$$r'_{out2} = r_{in} - e; \quad r''_{out2} = r_{in} + e.$$

З урахуванням очевидної умови $r_{out2} > r_{in}$ матимемо $r_{out2} = r_{in} + e = r_{out}$ ($\alpha + \alpha' = \pi$).

Подані розрахунки дають змогу визначити розташування отворів на решті сателітів ($j = \overline{2, K}$; K – кількість сателітів) за умови, що розташування отворів на першому ($j = 1$) відоме. Положення цівки в інерціальній системі координат визначається матрицею, що описує положення їх центрів у довільний момент часу:

$$\mathbf{M}_n(\varphi) = \begin{bmatrix} X_{M_n} \\ Y_{M_n} \end{bmatrix} = \begin{bmatrix} \cos(\varphi/u) & -\sin(\varphi/u) \\ \sin(\varphi/u) & \cos(\varphi/u) \end{bmatrix} \cdot \begin{bmatrix} -R \sin\left((n-1)\frac{2\pi}{N}\right) \\ R \cos\left((n-1)\frac{2\pi}{N}\right) \end{bmatrix},$$

де n – номер цівки ($n = \overline{1, N}$; N – кількість цівок).

Так само положення отворів у j -му сателіті можна записати як

$$\mathbf{B}_{jn}(\varphi) = \begin{bmatrix} X_{B_{jn}} \\ Y_{B_{jn}} \end{bmatrix} = \mathbf{M}_n(\varphi) + e \cdot \begin{bmatrix} \sin(\varphi + (j-1)\alpha) \\ \cos(\varphi + (j-1)\alpha) \end{bmatrix}.$$

Тоді точку E_{jn} (точку дотику цівки вихідної ланки з n -м отвором j -го сателіта) визначимо системою рівнянь:

$$\begin{cases} (X_{E_{jn}} - X_{M_n})^2 + (Y_{E_{jn}} - Y_{M_n})^2 = r_{in}^2; \\ (X_{E_{jn}} - X_{B_{jn}})^2 + (Y_{E_{jn}} - Y_{B_{jn}})^2 = r_{out}^2, \end{cases}$$

або в матричній формі

$$\begin{cases} \|\mathbf{E}_{jn} - \mathbf{M}_n\| = r_{in}; \\ \|\mathbf{E}_{jn} - \mathbf{B}_{jn}\| = r_{out}. \end{cases}$$

Знаючи положення точки E_{jn} , визначимо кут тиску в контактній парі “ n -а цівка вихідної ланки – відповідний отвір j -го сателіта” [9] – кут між вектором лінійної швидкості точки M_{jn} і вектором контактної сили, докладеної в точці E_{jn} (рис. 5):

$$\cos(\gamma_{jn}) = \frac{(\mathbf{M}_n - \mathbf{E}_{jn}) \cdot \frac{d\mathbf{M}_n}{d\varphi}}{\|\mathbf{M}_n - \mathbf{E}_{jn}\| \cdot \left\| \frac{d\mathbf{M}_n}{d\varphi} \right\|} = \cos\left(\frac{\pi}{2} + \varphi\left(1 + \frac{1}{u}\right) + \alpha(j-1) + \frac{2\pi}{N}(n-1)\right), \quad (6)$$

де $\mathbf{E}_{jn}(\varphi) = \begin{bmatrix} X_{E_{jn}} & Y_{E_{jn}} \end{bmatrix}^T$ – координати точки E_{jn} у системі OXY .

Кут тиску є визначальним для багатьох експлуатаційних показників КПП. Зокрема, від нього залежать значення сили в контактній парі (парі “цівка – отвір”) і, відповідно, контактні механічні напруження. Використовуючи отримані в роботі [9] результати, контактні сили в цих парах можна визначити за принципом можливих переміщень:

$$T_{вих} \delta\varphi_{вих} - \sum_{j=1}^K \sum_{n=1}^N F_{jn} \delta r_{jn} = 0, \quad (7)$$

де $T_{вих}$ – крутний момент на вихідній ланці редуктора; δr_{jn} – нормальне зміщення точки E_{jn} прикладання сили F_{jn} у n -й парі j -го сателіта за можливого повороту вихідної ланки на кут $\delta\varphi_{вих}$ (див. рис. 5).

У виразі (7) враховано всі ведучі контактні пари. Якщо за модулем кут тиску γ_{jn} , обчислений за формулою (6), більший за $\pi/2$, то пара є веденою, і, відповідно, для неї контактна сила дорівнює нулеві ($F_{jn}|_{|\gamma_{jn}| \geq \pi/2} = 0$). Для визначення решти значень контактних сил слід вивести залежність між контактною силою та зміною міжосьової відстані геометричних поверхонь контактних пар, яка загалом може бути нелінійною:

$$\begin{aligned} F_{jn} &\approx A \cdot (\Delta r_n)^v = A \cdot \left[\Delta\varphi \cdot \|\mathbf{E}_{jn}\| \cdot \cos(\theta_{jn}) \right]^v = \\ &= A' \cdot \left[\|\mathbf{E}_{jn}\| \cdot \cos(\theta_{jn}) \right]^v, \end{aligned} \quad (8)$$

де A й A' – деякі константи; $v = 1 + \delta$ – показник степеня ($\delta = 0_+$ – реальне значення δ може бути визначене експериментально).

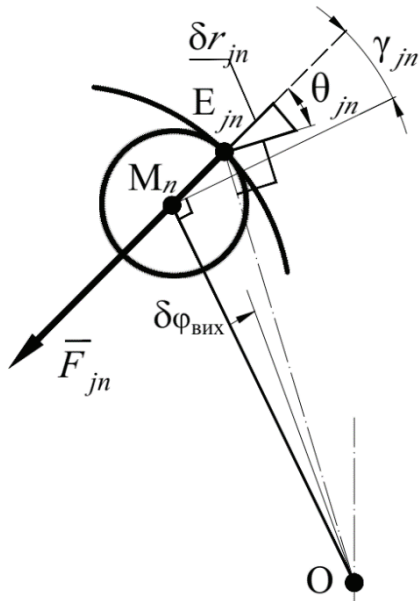


Рис. 5. Кут тиску та контактна сила в контактних парах

З трикутника OM_nE_{jn} (див. рис. 5) отримуємо вираз для кута θ_{jn} :

$$\cos(\theta_{jn}) = \frac{R}{\|E_{jn}\|} \cos(\gamma_{jn}).$$

Тоді вираз (8) для контактної сили набуде вигляду

$$F_{jn} = A' \cdot R^v \cdot \cos(\gamma_{jn})^v. \quad (9)$$

З урахуванням співвідношення $\delta r_{jn} = R \cdot \delta\varphi_{вих} \cdot \cos(\gamma_{jn})$ (див. рис. 5) і рівняння (9) вираз (7) матиме такий вигляд:

$$T_{вих} \delta\varphi_{вих} - A' \cdot R^{1+v} \cdot \delta\varphi_{вих} \cdot \sum_{(j,n)} \cos(\gamma_{jn})^{1+v} = 0,$$

звідки

$$F_{jn} = \frac{T_{вих}}{R} \cdot \frac{\cos(\gamma_{jn})^v}{\sum_{j=1}^K \sum_{n=1}^N \cos(\gamma_{jn})^{1+v}}. \quad (10)$$

Якщо припустити, що крутний момент на вхідній ланці редуктора постійний ($T_{вих} = T_{вх} \cdot u \cdot \eta$, де $T_{вх}$ – крутний момент на вхідному валу, η – ККД редуктора), то залежність контактної сили F_{jn} від кута повороту вхідного вала φ , як впливає з (10), визначатиметься функцією

$$f_{jn}(\varphi) = \frac{F_{jn}(\varphi)}{F_0} = \frac{\cos(\gamma_{jn})^v}{\sum_{j=1}^K \sum_{n=1}^N \cos(\gamma_{jn})^{1+v}}, \quad (11)$$

а її максимальне значення не перевищуватиме $F_0 = T_{вих}/R$ ($f_{jn}(\varphi) < 1$).

Слід зауважити, що у формулах (10) та (11) сума здійснюється лише для тих j і n , для яких виконується умова $\cos(\gamma_{jn}) \geq 0$, у якій кут тиску γ_{jn} обчислюється за формулою (6).

Також зазначимо, що для виконання умови збирання таких планетарних редукторів необхідно враховувати поворот отворів другого сателіта відносно першого. Цей кут можна обчислити за формулою $\psi_j = \min\left(j \cdot \frac{\alpha}{u} \pm k \cdot \frac{2\pi}{z}\right)$ ($k \in \mathbb{N}$).

Водночас цей кут не впливає на попередні розрахунки, оскільки поворот усіх отворів сателіта еквівалентний повороту зубів на сателіті.

Числові результати

Числові експерименти проведено для КПР із передатним числом $u = 20$, кількість отворів у сателітів (цівок вихідної ланки) якого дорівнює 3 ($N = 3$).

На рис. 6 зображено циклограми роботи контактних пар у редукторах із різним значенням кількості його сателітів K і кута між кривошипами α . З використанням цих циклограм нескладно обчислити “коефіцієнт перекриття контактних пар” [10]. На рис. 6, а, зокрема, показана циклограма для класичного виконання редуктора з $K = 2$ й $\alpha = 180^\circ$, отримана в роботі [9]. Як видно з рис. 6, а, у контакті постійно перебувають три пари. Зміна кута α спричинює зміну кількості одночасно активних контактних пар (рис. 6, б і 6, в), а отже й нерівномірність розподілу навантаження між ними. Так, за $\alpha = 120^\circ$ і $K = 2$ (див. рис. 6, б) кількість пар за один повний оберт кривошипа набуває двох можливих значень: 2 чи 4. Водночас за $\alpha = 90^\circ$ і незмінних K та N (див. рис. 6, в) кількість ведучих контактних пар у певний момент часу може дорівнювати 2, 3 чи 4. Останні два випадки свідчать про недоліки редуктора з $\alpha = 360^\circ/K$. Однак навіть виконання умови $\alpha = 360^\circ/K$ не є достатнім для вирівнювання розподілу навантаження між ведучими контактними парами за цикл, про що свідчить зображена на рис. 6, г циклограма – кількість ведучих контактних пар за $\alpha = 120^\circ$, $K = 3$ і $N = 3$ чергується між 3 та 6.

Виправити цей недолік можна збільшенням на один кількості цівок вихідної ланки. Так, за $\alpha = 120^\circ$; $K = 3$; $N = 4$ кількість ведучих контактних пар за цикл постійна та дорівнює 6 (рис. 6, д). За $N = 3$ постійна кількість ведучих контактних пар спостерігається лише для класичного вико-

нання планетарного редуктора ($K = 2$; $\alpha = 180^\circ$ (див. рис. 6, а)) та для подвоєної кількості сателітів ($K = 4$; $\alpha = 90^\circ$ (рис. 6, е)). Слід зауважити, що питання реалізації редуктора з $K = 4$ у цій роботі не порушено.

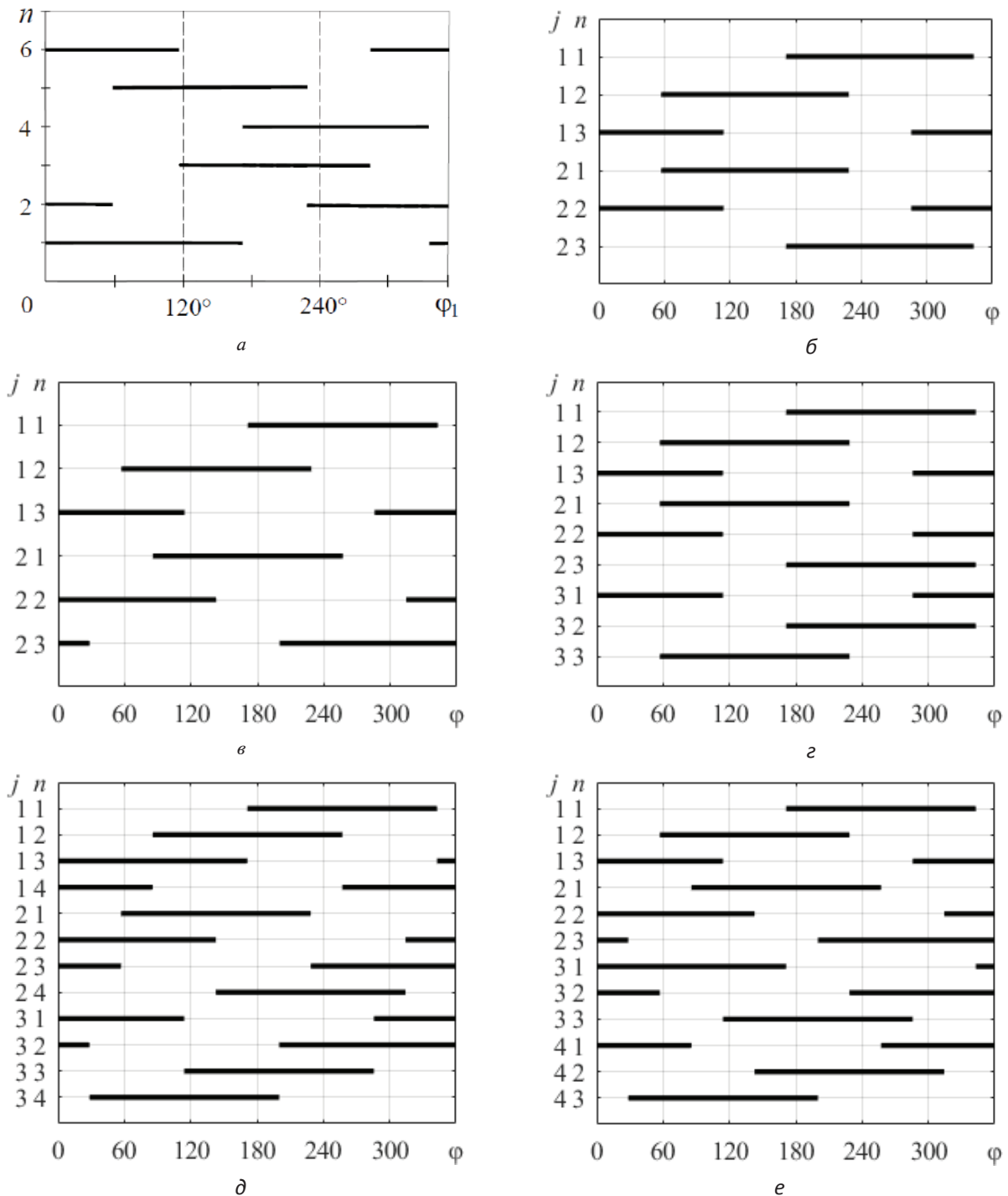


Рис. 6. Циклограми роботи ведучих контактних пар при різних геометричних конфігураціях редуктора: а – $\alpha = 180^\circ$; $K = 2$; б – $\alpha = 120^\circ$; $K = 2$; в – $\alpha = 90^\circ$; $K = 2$; г – $\alpha = 120^\circ$; $K = 3$; $N = 3$; д – $\alpha = 120^\circ$; $K = 3$; $N = 4$; е – $\alpha = 90^\circ$; $K = 4$; $N = 3$

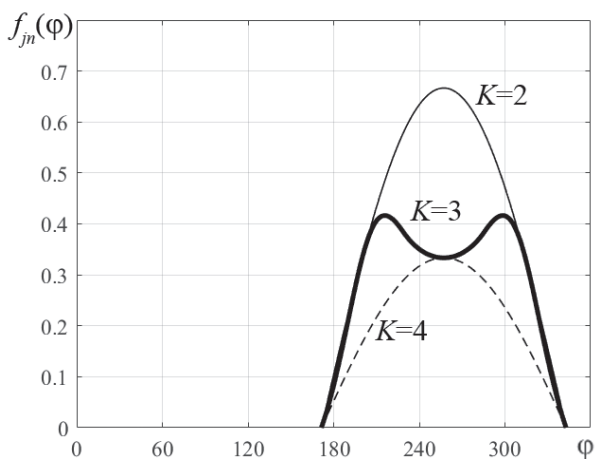


Рис. 7. Характер залежності контактної сили від кута повороту колінчастого вала

З аналізу поданих на рис. 6, а та 6, е циклограм впливає, що випадок ($K = 2$; $\alpha = 180^\circ$) є частинним випадком ($K = 4$; $\alpha = 90^\circ$), а кількість одночасно активних контактних пар залишається незмінною: 6 для $K = 4$ та 3 для $K = 2$. Нескладно передбачити, що збільшення кількості ведучих контактних пар забезпечує вдвічі менше максимальне значення контактної сили. Відповідні криві зображено на рис. 7 кривими $K = 4$ та $K = 2$. На рис. 7 показано також графік періодичної функції $f_{jn}(\varphi)$ (12), де φ – кут повороту колінчастого вала. Зміна кількості ведучих контактних пар призводить до складнішого характеру зміни активних у них контактних сил за цикл, які залежать від $f_{jn}(\varphi)$ – на рис. 7 крива $K = 3$ відповідає випадку $K = 3$; $\alpha = 120^\circ$.

Слід зауважити, що в розрахунках було прийнято лінійний характер залежності контактної сили від деформування матеріалу ($\delta = 0$). Однак це не применшує коректності результатів.

Значення крутного моменту на вихідному валу залежить як від крутного моменту на вході редуктора та його передатного числа, так і від ККД редуктора. Останній також залежить від тертя в контактних/кінематичних парах редуктора. Тому результати можуть бути використані під час уточненого розрахунку ККД КПП

і визначення нерівномірності крутного моменту на його вихідному валу.

Висновки

У роботі подано аналітичний метод побудови циклограм роботи контактних пар “отвір сателіта – цівка вихідної ланки” КПП залежно від його геометричних параметрів, зокрема, від кута між кривошипами колінчастого вала, кількості сателітів і цівок вихідної ланки. Вказана циклограма є визначальною під час як обчислення деталей редуктора на міцність, так і розрахунку експлуатаційних показників редуктора загалом. Слід зауважити, що отримане співвідношення для контактної сили як функції кута повороту колінчастого вала припускає нелінійну поведінку матеріалів деталей за його деформування. Водночас відповідний показник степеня має бути визначений експериментально.

На підставі отриманих співвідношень показано можливість реалізації редукторів із довільним значенням кута між кривошипами колінчастого вала α та кількості сателітів K і цівок вихідної ланки N . Однак, як показали розрахунки, умовою забезпечення рівномірності розподілу навантаження між контактними парами та збільшення “коефіцієнту перекриття” має бути дотримання співвідношення $\alpha = 360/K$. Достатня умова має бути перевірена чисельно через побудову циклограм роботи контактних пар. Водночас значення N може бути використане як допоміжний параметр для покращення умов роботи контактних пар.

За незмінної кількості одночасно активних ведучих контактних пар і, відповідно, рівномірної передачі навантаження між ними закон зміни контактної сили за цикл подібний до параболічного. В іншому випадку зміна контактної сили за цикл має досить складний характер.

Результати можуть бути використані для уточненого розрахунку ККД КПП, для визначення нерівномірності крутного моменту на його вихідному валу та під час ухвалення конструкторських рішень на етапі проектування КПП.

References

- [1] V.N. Kudryavtsev et al., *Planetary gears. Directory*. V.N. Kudryavtsev, and Y.N. Kirdyashev, Eds., Leningrad, USSR: Mashinostroyeniye, 1977.
- [2] S. Macovei, and I. Doroftei, “An overview on internal geared mechanisms with small difference between teeth number”, *IOP Conf. Ser.: Mater. Sci. And Eng.*, vol. 95, pp. 012053, Nov. 2015, doi: 10.1088/1757-899X/95/1/012053.
- [3] V. B. Tarabarina, and Z. I. Tarabarina, “Crank-planetary gearboxes with involute internal gearing with a difference in the numbers of teeth of the wheels $Z_d = 1$ ”, *Mechanics and Mechanical Engineering*, no. 1, pp. 1–12, 2006.

- [4] E. Chen, and D. Walton, "The optimum design of KHV planetary gears with small tooth differences", *Int. J. of Mach. Tools and Manufact.*, vol. 30, no. 1, pp. 99–109, Jan. 1990, doi: 10.1016/0890-6955(90)90045-K.
- [5] D. Yu, "KHV Planetary gearing", in *Gear technologies: Part I*, pp. 21–48, 1987, Part II, pp. 28–48, 1988.
- [6] W.-Y. Lin, K.-M. Hsiao, and J.-Y. Ke, "Optimum design of involute tooth profiles for K-H-V planetary drives with small teeth number differences", *J. of Advanced Mech. Design, Systems, and Manufacturing*, vol. 13, no. 1, p. jamdsm0007, 2019, doi: 10.1299/jamdsm.2019jamdsm0007.
- [7] Ch. Si, Zh. Chen, H. Zhang, Y. Zhang, and Ch. Cui, "Cosine gear planetary transmission with small teeth number difference", *IOP Conf. Ser.: Mater. Sci. and Eng.*, p. 012034, Mar. 2020, doi: 10.1088/1757-899X/793/1/012034.
- [8] M. A. Pavlovskiy, *Theoretical Mechanics: Textbook*. Kyiv, Ukraine: Tekhnika, 2004.
- [9] I. Yanchevsky, P. Komada, N. Stelmakh, and D. Lytvynenko, "Contact problem for crank-planetary reduction gear," *KPI Science News. Ser. Instrument making and information and measuring equipment*, no. 1, pp. 73–80, 2018, doi: 10.20535/1810-0546.2018.1.12116.8.
- [10] O. P. Zakhovaiko, *Theory of mechanisms and machines. Course of lectures*. Kyiv, Ukraine: NTUU "KPI", 2010.

I.V. Yanchevskiy, M.V. Niekhoroshykh

INFLUENCE OF GEOMETRIC PARAMETERS OF CRANKSHAFT-PLANETARY REDUCER ON CYCLOGRAPH OF WORK OF CONTACT PAIRS

Background. Crank-planetary reducers (CPR) have a number of advantages: compactness, high load capacity, high structural rigidity, to name a few. At the same time, there is no regulated methodology for their kinematic and force calculation, which limits their development and mass production. The creation of effective methods for calculating these reducers to improve their performance at the design stage is of great practical importance.

Objective. The aim of this work is to develop an analytical method for calculating the angle of pressure and the contact force in contact pairs "satellite hole – pin of output link" of the CPR depending on the angle between the cranks of its crankshaft (V-shaped), number of satellites and number of pins. Based on the developed method, check the possibility of designing a CPR with arbitrary values of these parameters and investigate their impact on the cyclogram of contact pairs and the associated maximum value of contact force, taking into account the nonlinearity of materials' deformation.

Methods. At the first stage, based on the formulas of analytical geometry, the possibility of theoretical implementation of CPR with an arbitrary value of the angle between the crankshaft cranks, the number of satellites and the number of pins of the output link is shown. The values of pressure angles in the contact pairs "satellite hole – pin of output link" are analysed and on the basis of the received results the cyclogram of work of leading pairs is calculated. In the second stage, using the principle of possible displacements, the formula for calculating the contact force in the leading contact pairs is written as a function of the input shaft rotation angle. When calculating the contact force, a nonlinear model of material deformation was used.

Results. An analytical method for constructing a cyclogram of the contact pairs of a CPR and the contact forces calculation was developed depending on the geometric parameters of the reducer. The possibility of theoretical implementation of reducers with arbitrary value of the angle between the crankshaft cranks, the number of satellites and the number of output link's pins is shown

Conclusions. The dependence of the contact force is obtained as a rotation angle function, the number of satellites and the number of source pins. With a constant number of simultaneously active leading contact pairs and, accordingly, a uniform load transfer between them, the law of change in the contact force per cycle is similar to parabolic. Otherwise, the change in the contact force per cycle is quite complex. The obtained results can be used in the refined calculation of the efficiency of the CPR and in making various design decisions at the design stage.

Keywords: crank-planetary reducer; geometric parameters; cyclogram; contact force; pressure angle.

Рекомендована Радою
Навчально-наукового механіко-машинобудівного Інституту
КПІ ім. Ігоря Сікорського

Надійшла до редакції
26 жовтня 2021 року

Прийнята до публікації
14 лютого 2022 року

DOI: 10.20535/kpissn.2021.4.239236

УДК 620.92

Ю. М. Мацевитий¹, Я. М. Буштець^{2*}¹Інститут проблем машинобудування ім. А. М. Підгорного НАН України, Харків, Україна²ТОВ “Глобал Білгі”, Харків, Україна

*corresponding author: bushtetsyana@gmail.com

РІЗНОВИДИ ТА ТЕХНІЧНІ ХАРАКТЕРИСТИКИ СОНЯЧНИХ КОЛЕКТОРІВ

Проблематика. Енергетичний ресурс є одним із головних складників успішності країни, а можливість використання екологічно чистої та доступної відновлюваної енергії сонячного випромінювання є важливим завданням. Нині існує велике різноманіття геліосистем за технічним рішенням і сферами використання: термо- та фотоелектричні, сонячні станції для промислового чи приватного видобування електроенергії, гарячого водопостачання, опалення, кондиціонування повітря тощо. Але питання щодо ефективної інсоляції та вибору матеріалів для цих систем відкрите. Тому є актуальним розроблення дієвих сонячних колекторів (СК) з раціональними геометричними та теплофізичними параметрами.

Мета дослідження. Знайти та розробити СК з максимальною пропускною здатністю абсорбера та мінімізованими тепловтратами.

Методика реалізації. Аналіз патентів СК за їхньою формою та технічними характеристиками, розробка нової моделі СК.

Результати дослідження. Згідно з проведеним оглядом, нові розробки геліоколекторів відзначаються оригінальністю та новизною технічних рішень. Як матеріал абсорбера використовують гофровану трубу, сотовий полікарбонат, активоване вугілля, плутанку теплопровідного дроту; застосовують додаткове теплоакмуляційне покриття з розчином глауберової солі. Винаходи можна застосовувати в системах опалення та гарячого водопостачання, підігріву повітря приміщень і в обладнанні, що перетворює електромагнітне випромінювання сонця на теплову енергію для нагрівання рідкого теплоносія, а також на електроенергію. Для досягнення максимальної потужності колектора використовують приймальну поверхню з покриттям із високою пропускною здатністю сонячної енергії, а для мінімізації тепловтрат додатково встановлюють теплоізоляцію. Розроблено та запатентовано принципово новий за геометричною формою сферичний СК.

Висновки. Розробка нових й удосконалення наявних СК забезпечує енергонезалежність як у приватному секторі, так і в масштабах країни, а також дає змогу дбати про довкілля. Розроблений сферичний СК однаково ефективно поглинає променисту енергію сонця протягом дня.

Ключові слова: використання сонячної енергії; сонячний колектор; абсорбер сонячного колектора; пропускна здатність абсорбера.

Вступ

Сонячний колектор – це пристрій для збору енергії випромінювання Сонця. З часів Ораса Бенедикта де Соссюра, швейцарського винахідника, автора прототипу СК – геліотермометра (сонячної печі) 1780 року – принцип дії СК майже не змінився. Але інтерес дослідників до розробки приладу з максимальним поглинанням та мінімальними втратами є достатньо високим – це показує вагома світова база патентів.

Учені працюють над матеріалом, формою та технічним призначенням СК.

Основою СК є поглинальний елемент – абсорбер, який з'єднаний з теплопровідною системою. За типом абсорбера СК поділяють на плоскі, вакуумні, коаксіальні та пір'яні.

Плоский СК – це металева пластина з поглинальним покриттям, подібним до “меандру” чи “арфи”, по якій циркулює теплоносій. Абсорбер укладають у корпус із прозорою лицьовою стінкою, через яку проникає сонячне випромінювання.

Рекомендуємо цитувати цю статтю так: Ю. М. Мацевитий, Я. М. Буштець, “Різновиди та технічні характеристики сонячних колекторів”, *Наукові вісті КІП*, № 4, с. 92–98, 2021. doi: 10.20535/kpissn.2021.4.239236.

Please cite this article as: Y. M. Matsevytyi, and Y. M. Bushtets, “Varieties and technical characteristics of solar collectors”, *KPI Science News*, no. 4, pp. 92–98, 2021. doi: 10.20535/kpissn.2021.4.239236.

нювання, а тильну стінку оснащують утеплювачем.

Вакуумний СК – це ряд паралельних прозорих трубчастих профілів. Вакуумна труба нагадує “термос” (у трубу з більшим діаметром вставляють трубу з меншим діаметром, між якими відкачують повітря). Завдяки вакууму вдається отримати майже повну теплоізоляцію. Внутрішню трубу встеляють високоселективним покриттям, що дає можливість виробляти енергію за мінусової температуридовкілля.

У коаксіальному СК у прозору скляну колбу вкладають мідну трубку-гільзу. Всередині трубки є рідина, що має низьку температуру кипіння, наприклад, вода під низьким тиском. Під впливом сонячних променів рідина (вода) нагрівається та закипає, а пара підіймається та надходить до теплообмінника, в якому нагріває теплоносії. Потім відбувається конденсація охолодженої пари на стінках теплообмінника. За потреби процес можна повторити.

Пір’яний вакуумний СК – це одностінна скляна колба, в якій товщина стінок більша, ніж у коаксіального СК. Розташований усередині колби абсорбційний елемент, який виготовляють із міді, додатково по всій довжині посилюють гофрованою пластиною (за формою нагадує перо), яка має високорівневе енергопоглинальне напilenня. Така конструктивна особливість дає змогу отримати вакуум безпосередньо в тепловому каналі.

Постановка задачі

Метою дослідження є знайти та розробити ефективні СК з максимальною пропускною здатністю абсорбера та мінімізованими тепловтратами. Навести світові розробки СК за їхньою формою та технічними характеристиками.

Аналіз конструкцій СК

За останні десять років патентна база України поповнилася цікавими технічними розробками СК.

Винахідники В. А. Істратов і Д. М. Раус запропонували конструкцію СК (рис. 1), що містить прозорий корпус напівсферичної форми 1 й абсорбер – чорну гофровану трубу 2, заповнену рідиною, фізико-хімічні властивості якої відповідають матеріалу труби щодо антиагресивності. Гофрована труба 2 розташована в корпусі спіралеподібно, стискаючись із внутрішньою поверхнею півсфери [1].

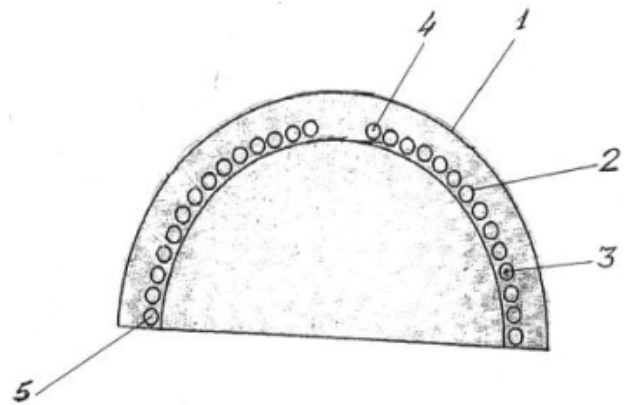


Рис. 1. Сонячний колектор (патент UA 73587 від 25.09.2012): 1 – світлопрозорий напівсферичний корпус; 2 – гофрована труба; 3 – теплоносії; 4, 5 – торці

Сонячні промені, потрапляючи на СК, проходять світлопрозорий напівсферичний корпус 1, виконаний із сотового полікарбонату з високою світлопроникністю, потрапляють на зовнішню поверхню абсорбера – чорної гофрованої труби 2, через стінку якої проходять до теплоносія 3 (води). Вода починає рух у трубі з торця 4, нагрівається та надходить із торця 5 до споживача. Модель можна використовувати в системах сонячного опалення та гарячого водопостачання.

Цікаву конструкцію СК (рис. 2) запропонували вчені з Українського державного хіміко-технологічного університету під головуванням К. М. Сухого. Тут поглиначем сонячної енергії є добре змочене активоване вугілля 2, що дає можливість ефективно використовувати природну циркуляцію теплоносія завдяки капілярним силам [2].

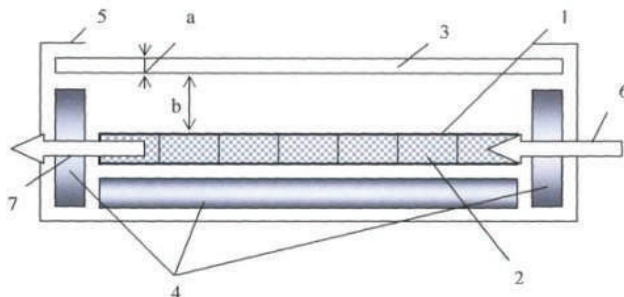


Рис. 2. Сонячний колектор (патент UA 53855 від 25.10.2010): 1 – абсорбер; 2 – активоване вугілля; 3 – світлопрозоре огороження; 6, 7 – торці

Принцип колектора роботи такий: сонячні промені через світлопрозоре огороження 3

потрапляють на абсорбер 1, виготовлений із сотового полікарбонату, в якому активоване вугілля 2 поглинає сонячну енергію та передає як теплову енергію до теплоносія. З торця 6 теплоносія входить в абсорбер колектора 1, нагрівається та виходить із торця 7.

Оригінальною є й пропозиція від колег із Національного авіаційного університету. Запропонований СК (рис. 3) містить систему паралельних скляних колб 3 з чорною трубою 1 всередині, в якій протікає рідина типу антифризу. До чорної труби 1 на виході скляної колби 3 підключений теплопередавач 5 – теплоізолювана труба – та помпа 4, яка перекачує гарячий антифриз у теплообмінник, до якого можна підвести гаряче водопостачання та центральне опалення. Сонячні батареї 6, установлені під нахилом до паралельно розташованих скляних вакуумних колб 3, віддзеркалюють сонячні промені задля нагрівання колб 3. У торцях розташовані з'єднані з сонячними батареями додаткові поверхні 7, що віддзеркалюють, змінюючи нахил за допомогою електромоторів EM1 і EM2 та приводу управління електромоторами, який містить таймер і вимірювач інтенсивності сонячного випромінювання [3].

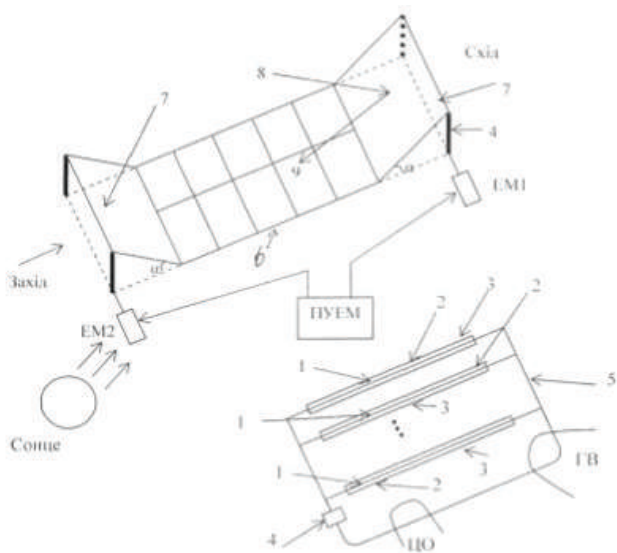


Рис. 3. Сонячний колектор із фокусувальними управляючими дзеркалами (патент UA 99699 від 25.06.2015): 1 – труба; 2 – вакуум; 3 – скляні колби; 4 – помпа; 5 – теплопередавач; 6 – сонячні батареї, 7 – віддзеркалювальні поверхні

Група винахідників з Волинської області СК (рис. 4), який складається з вакуумної трубки 1 і теплової трубки 2, оберненої плутанкою

теплопровідного дроту 3 з енергопоглинальним покриттям. Сонячні промені проходять через стінку скляної трубки 1 і потрапляють на енергопоглинальне покриття дроту 3, перетворюючись на тепло, яке переходить безпосередньо до теплової трубки. Відбивання світла на поверхнях ділянок плутанки теплопровідного дроту 3 створює тепловий ефект [4]. Сонячний колектор можна використовувати в пристроях, що переробляють електромагнітне випромінювання сонця в теплову енергію для нагрівання рідкого теплоносія, а також в електричну енергію.

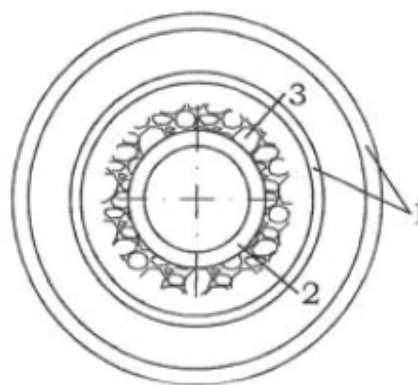


Рис. 4. Переріз трубки СК (патент UA 117717 від 10.07.2017): 1 – вакуумна трубка; 2 – тепла трубка; 3 – плутанка теплопровідного дроту

Учені з Національного університету “Львівська політехніка” запропонували для обігрівання житлових приміщень систему повітряного опалення [5] (рис. 5), яка містить плоский сонячний колектор 1, з’єднаний гнучкою вставкою 2 з повітропроводом 3, на вході якого встановлено вентилятор 5. Із зовнішнього боку повітропровід 3 утеплений шаром теплоізоляційного матеріалу 4, а з внутрішнього – теплоакмуляційним покриттям 6 – матрицею з шаром теплоізоляційного матеріалу 11 із насиченим розчином глауберової солі 10 – і знизу обмежений підвісною стелею 7 із повітророзподільними отворами 8 і регульовальними заслінками 9. Система повітряного опалення працює так: нагріте в плоскому сонячному колекторі 1 повітря через гнучку вставку 2 надходить за допомогою вентилятора 5 у повітропровід 3. Далі повітря проходить через теплоакмуляційне покриття 6 і через повітророзподільний отвір 8 надходить у приміщення, яке обігрівають. Якщо обігрівати не потрібно, регульовальну заслінку 9 закривають, а тепло накопичується теплоакмуляційним покриттям 6.

За температури повітря довкілля вище за 32 °С розчин глауберової солі перетворюється на рідину, а за температури до 20 °С і нижче відбувається фазовий перехід речовини з вивільненням тепла, внаслідок якого розчин кристалізується.

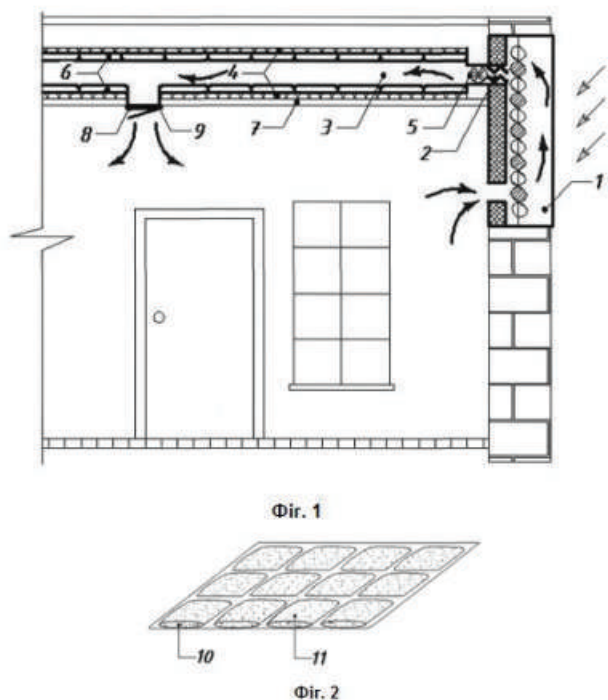


Рис. 5. Система повітряного опалення (а) та матриця з насиченим розчином глауберової солі (б) (патент UA 113039 від 10.01.2017): 1 – плоский сонячний колектор; 2 – гнучка вставка; 3 – повітропровід; 4 – шар теплоізоляційного матеріалу з зовнішнього боку; 5 – вентилятор; 6 – теплоаккумуляційне покриття; 7 – підвісна стеля; 8 – повітророзподільні отвори; 9 – регульовальні заслінки; 10 – насичений розчин глауберової солі; 11 – шар теплоізоляційного матеріалу з внутрішнього боку

Винахідник із Греції запропонував СК (рис. 6), який складається з теплопровідної пластини 3, прикріпленої до задньої частини поверхні сонячного теплового поглиначя 1 і встановленої на трубах 2 з теплоносієм, що поступає до будь-якого СК, збільшуючи індукційну поверхню та площу теплопокриття [6].

Розробник із Канади запропонував конструкцію [7], сегмент якої показано на рис. 7. Сонячні промені, відбиваючись від параболічної стінки 2, потрапляють на трубу 3 з теплоносієм. Наведену конструкцію можна обладнати додатково системою обертання навколо горизонтальної осі 4 та підсилити встановленням внутрішньої труби 3.

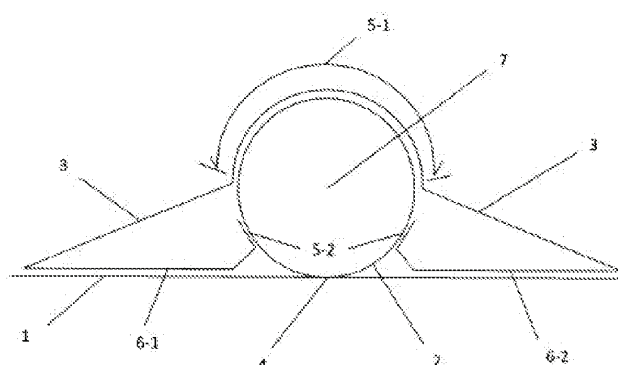


Рис. 6. Сонячний колектор (патент GR 20190100215 від 15.12.2020): 1 – сонячний тепловий поглинач; 2 – труби; 3 – теплопровідна пластина.

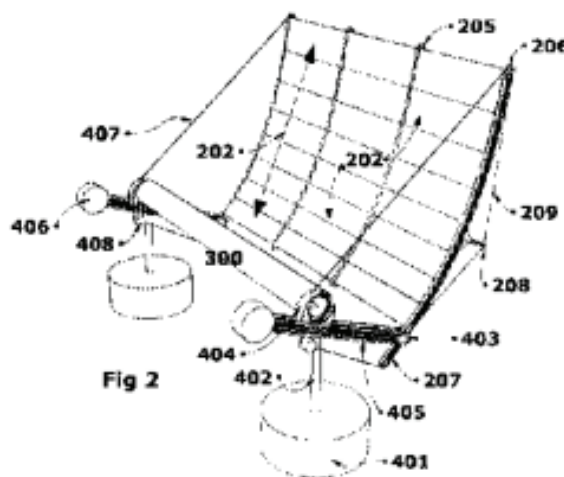


Рис. 7. СК (патент WO 2021119795 від 24.06.2021): 1; 2 – параболічна стінка; 3 – труба з теплоносієм; 4 – система обертання навколо горизонтальної осі

Винахідники з Казахстану запропонували таке технічне рішення (рис. 8). Сонячну енергію поглинає геліоколектор 1, вона проходить через ізоляційний прозорий склопакет 2, нагріває в трубах 3 воду, яку видаляють із колектора, а на її місце надходить холодна вода з трубопроводу з вентилям для холодної води 8, а з сифона 7 бака-дозатора 9 відбувається постійна термосифонна циркуляція за допомогою циркуляційної труби 10. Далі рідина потрапляє в тепловупомпу 11, яка складається з випарника 12 конденсатора, в якому теплообмінник виконаний як спіраль, поглинаючи тепло теплоносія, опускає його температуру нижче температури атмосферного повітря за допомогою дросельного клапана 14, сприяючи додатковому поглинанню тепла з атмосферного повітря.

У тепловій pompі енергія теплоносія з відносно низькою температурою переходить до теплоносія теплообмінника конденсатора 15 як спіралі з вищою температурою, що збільшує інтенсивність теплообміну. Для здійснення такого циклу використовують компресор 13 з електроприводом 17. Далі за допомогою теплообмінника конденсатора 15 тепло від теплової помпи передають у бак-акумулятор теплообмінника системи опалення 18. Установка має два контури з автоматичними циркуляційними помпами 19 і 20 для циркуляції рідини між геліоколекторами та випарником, а також між конденсатором і бак-акумулятором. Температуру води доводять до необхідного технологічного рівня та подають її до споживача для гарячого водопостачання й опалення [8].

В Інституті проблем машинобудування ім. А. М. Підгорного нами розроблено та запатентовано [9] принципово новий за геометричною формою сферичний СК (рис. 9, а). Він містить теплоприймач 1, виконаний із навитої сферичною гвинтовою лінією єдиної металевої трубки, на яку нанесено селективне світлопоглинальне покриття. Теплоприймач 1 розміщено всередині вакуумованої скляної сфери 2, що складається з двох півсфер, герметично встановлених у профільній прокладці 3 з вакуумної гуми. Підвідний і відвідний патрубки теплоприймача

1 герметично ущільнені у втулках 4, також виконаних із вакуумної гуми. Усередині профільної прокладки 3 (рис. 9, б) вбудовано опукле дзеркало 5 як кульовий пояс із радіусом R , рівним радіусу внутрішньої сфери теплоприймача, та висотою, рівною висоті профільної прокладки 3. З внутрішньої порожнини сфери 2 відкачано повітря зі ступенем вакууму, якого зазвичай досягають у промисловості. Сфера 2 з розміщеним усередині теплоприймачем 1, профільною прокладкою 3 та втулками 4 встановлена вертикально у двоплечовому кронштейні 6.

На вхід у підвідний патрубок надходить (показано стрілкою) холодний рідкий теплоносії (вода), що рухається гвинтовою лінією теплоприймача 1 і виходить через відвідний патрубок. Під час руху теплоносії постійно нагрівається променистою сонячною енергією, що впливає через прозору сферу 2 на селективне світлопоглинальне покриття трубчастого теплоприймача 1. Сонячні промені протягом світлового дня, проникаючи між витками теплоприймача 1, потрапляють на різні ділянки ввігнутого назовні сферичного дзеркала 5 і відбиваються у напрямку внутрішніх затієних ділянок витків теплоприймача 1. Вакуум у прозорій сфері практично повністю виключає тепловтрати в теплоприймачі 1.

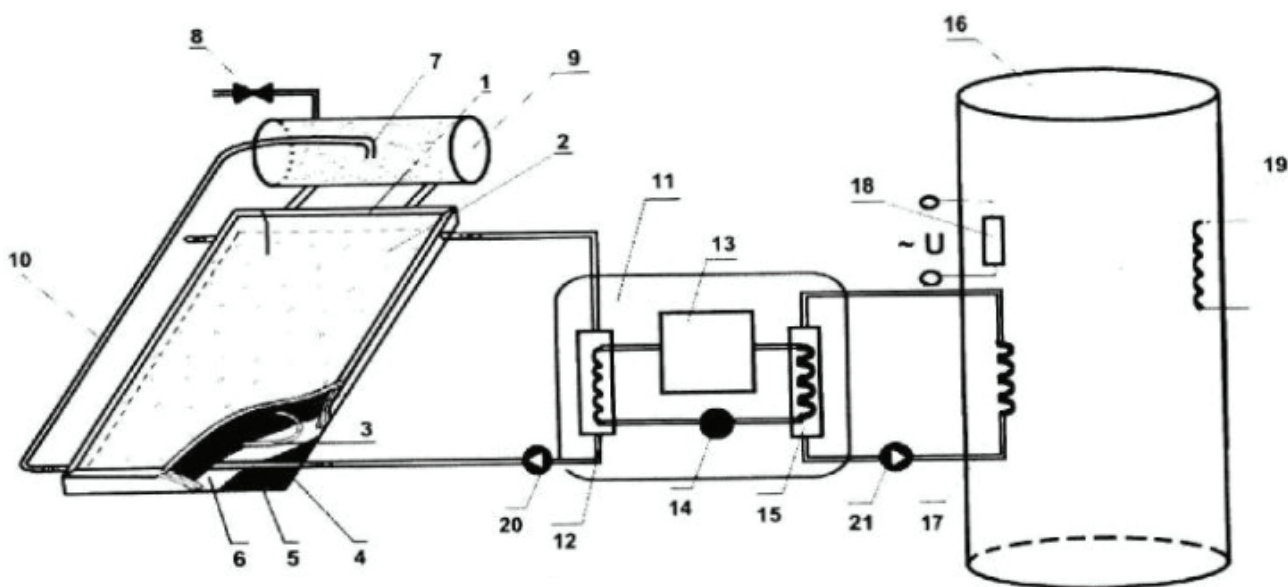


Рис. 8. СК (патент KZ 33741 від 02.07.2019): 1 – геліоколектор; 2 – ізоляційний прозорий склопакет; 3 – труби; 4; 5; 6; 7 – сифон; 8 – вентиль для холодної води; 9 – бак-дозатор; 10 – циркуляційна труба; 11 – тепла помпа; 12 – випарник; 13 – компресор; 14 – дросельний клапан; 15 – теплообмінник конденсатора; 16; 17 – електропривід; 18 – бак-акумулятор теплообмінника системи опалення; 19, 20 – автоматичні циркуляційні помпи

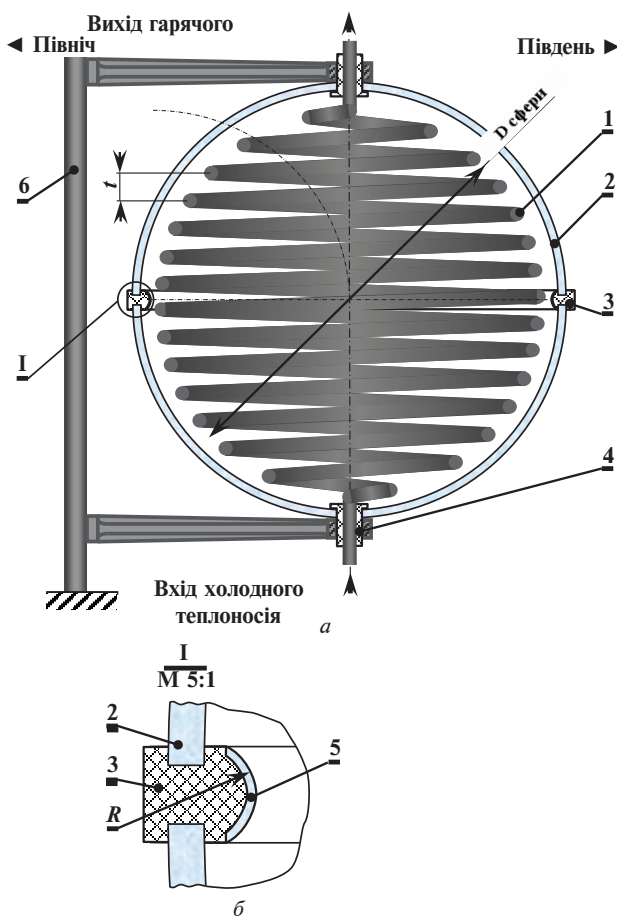


Рис. 9. Конструкція сферичного СК (а) та профільна прокладка (б): R – радіус; 1 – теплоприймач; 2 – вакуумована скляна сфера; 3 – профільна прокладка; 4 – втулки; 5 – опукле дзеркало; б – двоплечовий кронштейн

Далі гарячий теплоносій надходить для споживання (бак-акумулятор, опалення, гаряче водопостачання тощо). Вертикальне розташування входу та виходу трубчастого теплоприймача 1 у двоплечовому кронштейні 6 дає змогу знизити

гідроопір рідкого теплоносія, що проходить через нього.

Оскільки кривизна скляної сфери запропонованого колектора в десятки разів менша за кривизну скляних трубок плоского СК, знижується відбивна та підвищується пропускна здатність сонячних променів. Сферична форма не вимагає будь-яких додаткових механізмів, щоб слідкувати за рухом сонця. Завдяки такій формі колектор постійно «стежить» за ним, тобто фактично однаково сприймає променисту енергію сонця, коли воно рухається небосхилом на різній висоті та під різними кутами азимута. Площа під установлення сферичного колектора майже вчетверо менша за площу під монтаж плоского колектора.

Висновки

Згідно з оглядом, нові розробки геліоколекторів відзначаються оригінальністю та новизною технічного рішення. Як матеріал абсорбера застосовують гофровану трубу, сотовий полікарбонат, активоване вугілля, плутанку теплопровідного дроту, додаткове теплоаккумуляційне покриття з розчином глауберової солі. Винаходи можна застосовувати в системах опалення та гарячого водопостачання, в підігріві повітря приміщень і в обладнанні, що перетворює електромагнітне випромінювання сонця на теплову енергію для нагріву рідкого теплоносія, а також – на електричну енергію.

Авторами статті створено та запатентовано сферичний СК. Сферична форма дає змогу ефективно використовувати його протягом дня. Завдяки такій конструкції колектор однаково сприймає променисту енергію сонця, коли воно «рухається» небосхилом на різній висоті та під різними кутами азимута.

References

- [1] Solar collector, by V. A. Istratov, and D. M. Raus. (2012, Sep. 25). UA Patent 73587 [Online]. Available: <https://base.uipv.org/searchINV/>
- [2] Solar collector, by M. P. Sukhyi, Y. M. Kozlov, and K. M. Sukhyi. (2010, Oct. 25). UA Patent 53855 [Online]. Available: <https://base.uipv.org/searchINV/>
- [3] Solar collector with focusing control mirrors, by V. M. Syniehlazov et al.. (2015, Jun. 25). UA Patent 99699 [Online]. Available: <https://base.uipv.org/searchINV/>
- [4] Solar collector, by S. A. Mironov, V. I. Reshetnikov, A. A. Mironov, and V. V. Polovets. (2017, Jul. 10). UA Patent 117717 [Online]. Available: <https://base.uipv.org/searchINV/>
- [5] Air heating system, by V. M. Zhelykh, and Kh. R. Kozak. (2017, Jan.10). UA Patent 113039 [Online]. Available: <https://base.uipv.org/searchINV/>

- [6] Lamella for solar collector's heat transfer, by F. M. Athanasiou. (2020, Dec. 15). GR Patent 20190100215 [Online]. Available: https://patentscope.wipo.int/search/en/detail.jsf?docId=GR317308168&_cid=P22-KR945M-88396-2
- [7] Solar energy collector, by R. De Chazal. (2021, Jun. 24). WO Patent 2021119795 [Online]. Available: https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2021119795&_cid=P22-KR94A3-89006-1
- [8] Dual-circuit solar unit with thermosiphon circulation, by E. N. Amerhaliiev, M. M. Kunelbaev, O. A. Auelbekov, N. S. Kataiev, A. U. Kalizhanova, and A. Kh. Kozbakova. (2019, Jul. 2). KZ Patent 33741 [Online]. Available: https://patentscope.wipo.int/search/en/detail.jsf?docId=KZ320873366&_cid=P22-KR94ED-89540-1
- [9] Spherical solar collector, by A. I. Tsentsyper, M. O. Safonov, S. F. Lushpenko, and Y. M. Bushtets. (2014, Apr. 10). UA Patent 105112 [Online]. Available: <https://base.uipv.org/searchINV/>

Yu. M. Matsevytyi, Ya. M. Bushtets

VARIETIES AND TECHNICAL CHARACTERISTICS OF SOLAR COLLECTORS

Background. Energy resource is one of the main components of a successful country. The possibility of using environmentally clean and affordable renewable energy and creating systems for taking solar radiation is a priority in ensuring the economic energy independence of the state. At present, there are a variety of solar systems according to the technical solution and their use: thermo- and photovoltaic, systems with metal absorbers, solar stations for industrial or private electricity generation, for hot water supply, heating, air conditioning, and other technological needs. But questions about effective insolation and the choice of materials for these systems remain open. Therefore, the development of efficient solar collectors (SC) with a rational choice of their geometric and thermophysical parameters is absolutely relevant.

Objective. The purpose of the paper is to explore and develop efficient devices with maximum absorber throughput and minimization of heat loss.

Methods. Analysis of developments of solar collectors by their shape and technical characteristics.

Results. According to the review, new developments of solar collectors are marked by the originality and novelty of the technical solution. Corrugated pipe, cellular polycarbonate, activated carbon, a tangle of heat-conducting wire are used as the absorber material, and the additional heat-accumulating coating – a solution of Glauber's salt is applied. The inventions can be used in heating and hot water supply systems, indoor air heating, and in equipment that converts electromagnetic radiation from the sun into thermal energy for heating liquid coolants, as well as into electrical energy. For the maximum power of the collector, a receiving surface with a coating of high solar energy capacity is used and thermal insulation is additionally installed to minimize heat loss. The authors of the article developed and patented a fundamentally new geometric spherical SC.

Conclusions. The creation of new and improvement of existing SC ensures energy independence both in the private sector and nationwide, as well as to take care of the environment. The spherical shape allows you to use it effectively during the day.

Keywords: solar energy utilization; solar collector; solar collector absorber; throughput of the absorber.

Рекомендована Радою
Інституту проблем машинобудування
ім. А.М. Підгорного НАНУ

Надійшла до редакції
26 серпня 2021 року

Прийнята до публікації
14 лютого 2022 року

DOI: 10.20535/kpissn.2021.4.252713

UDC 620.179.14

Grigoriy S. Tymchik, Oleksandr O. Podolian, Maksym S. Druzhev
Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine

CONTROLLING THE THICKNESS OF THE PLASTIC LAYER APPLIED USING 3D PRINTING BY ELECTROMAGNETIC ACOUSTIC METHOD

Background: Currently, the EMA method is not used to control the thickness of the layer of plastic applied to a metal platform using 3D printing, which significantly limits the scope of its use. However, EMA flaw detectors are widely used to control the quality of the metal platform itself. The solution of the problem of radiation research will allow carrying out complex control of the thickness of the layer of plastic applied by 3D printing, which will increase the efficiency of ultrasonic flaw detection in the reliability and speed of work.

Objective: This article considers the possibility of determining the thickness of the layer applied by 3D printing using electromagnetic-acoustic (EMA) method.

Methods: The analysis of the relationship between the thicknesses of the layer of plastic applied by 3D printing on the acoustic pressure created in it during the control using EMA transducers. The influence of the thickness of the plastic layer applied to the metal platform by means of 3D printing on the acoustic parameters of the EMA transducer was investigated with the help of mathematical modeling. Experimentally obtained dependences showing the influence of the thickness of the layer applied by 3D printing on the generated acoustic pressure.

Results. The effect of the thickness of the plastic layer applied to a metal platform using 3D printing on the acoustic parameters of the EMA transducer is studied using mathematical modelling.

Dependences showing the effect of the thickness of the plastic layer applied to the metal platform using 3D printing on the created acoustic pressure are obtained experimentally. The maximum acoustic pressure is created when there is no thickness of the plastic layer applied to the metal platform using 3D printing. The pressure drops sharply, as the layer thickness increases.

Conclusions: Good convergence of the outcome of theoretical and experimental studies is demonstrated, with the approximation error of experimentally obtained data generally not exceeding 5%.

The study is based on a widely proven approach to analyzing the process of acoustic wave formation by an EMA transducer. The reliability of the results obtained is confirmed by the correct use of the mathematical tools and good convergence of the outcome of theoretical and experimental studies.

Keywords: EMA, converter, 3d, Printing, layer, plastic, thickness, acoustic pressure, non-destructive testing.

Introduction

Among the studies related to the development of non-destructive testing equipment, the search for non-contact methods for excitation and recording ultrasound in solids is the most important one [1]. Progress in this area has been achieved through the use of the electromagnetic acoustic (EMA) method of excitation and reception of ultrasonic vibrations.

Nowadays, the EMA method is not used to control the thickness of the plastic layer applied to a metal platform using 3D printing, which signifi-

cantly limits the scope of its application. Nevertheless, EMA flaw detectors are widely used to control the metal platform's quality. The problem of controlling the thickness of the plastic layer applied by 3D printing using EMA transducer can be addressed only in joint study of radiation, magnetic field [2, 3] and probe pulses formation [4]. The solution of the radiation study's problem is to allow for comprehensive control of the thickness of the plastic layer applied using 3D printing, which will increase the efficiency of ultrasonic flaw detection in terms of reliability and speed of work.

Рекомендуємо цитувати цю статтю так: Г.С.Тимчик, О.О.Подольян, М.С. Друзев, «Контроль товщини шару пластику нанесеного за допомогою 3Д друку електромагнітно-акустичним методом», *Наукові вісти КПІ*, № 4, с. 99–105, 2021. doi: 10.20535/kpissn.2021.4.252713.

Please cite this article as: Grigoriy S. Tymchik, Oleksandr O. Podolian, Maksym S. Druzhev, “Controlling the Thickness of the Plastic Layer Applied Using 3D Printing by Electromagnetic Acoustic Method”, *KPI Science News*, no. 4, pp. 99–105, 2021. doi: 10.20535/kpissn.2021.4.252713.

Problem statement

The problem is to analyze the relationship between the thickness of the plastic layer applied using 3D printing and the acoustic pressure created therein during controlling with the use of EMA transducers.

Analysis of the dependence of acoustic wave parameters on the thickness of the plastic layer applied to a metal platform using 3D printing

Based on the studies of a number of authors [5,6,11], in our case, a wire with simple harmonic current is mounted on the surface of a layer of plastic applied using 3D printing on a metal platform [5,11]:

$$I = I_0 \cdot \cos(\omega \cdot t) = I_0 \cdot \cos(2 \cdot \pi \cdot f \cdot t)$$

Where I_0 – is current amplitude value, ω – is angular frequency, f – is frequency of current in the wire.

As a rule, external magnetic fields are used to improve the quality of EMA transduction. There is an explanation in the ferromagnetism theory, according to which eddy currents that occur due to domain displacement have the greatest impact on the attenuation of ultrasonic waves in the object. The external magnetic field reduces ultrasound attenuation by or-

dering the domain structure. Consequently, the total magnetic field can be expressed as follows[6,11,12]:

$$B_{\Sigma} = B_{\approx} + B_{-}$$

Thus, we can put down the equation of acoustic pressure's distribution on the surface [6,11]:

$$p = -\mu_0 \cdot \mu \cdot \frac{I_0 \cdot h^2}{4 \cdot \pi \cdot (h^2 + y^2)^2} \cdot (1 + \cos(4 \cdot \pi \cdot f \cdot t)) - \frac{I_0 \cdot h \cdot B_{-}}{\pi \cdot (h^2 \cdot y^2)} \cdot \cos(2 \cdot \pi \cdot f \cdot t)$$

Where B_{-} – is induction of the external constant magnetic field, h – is thickness of the plastic layer applied using 3D printing, y – is distance from the wire projection along the surface of the object under test, $\mu_0 = 4 \cdot \pi \cdot 10^{-7} \text{Gn} / \text{m}$ – is magnetic constant, μ – is magnetic permeability of the controlled material.

Above equation demonstrates that there is an inverse relationship between the thickness of the plastic layer applied using 3D printing h and acoustic pressure. That is when increasing h the acoustic pressure will decrease.

As a result of modelling, the relationship between the thickness of the layer applied using 3D printing and the acoustic pressure was established, which is shown in the following diagram, Fig.1:

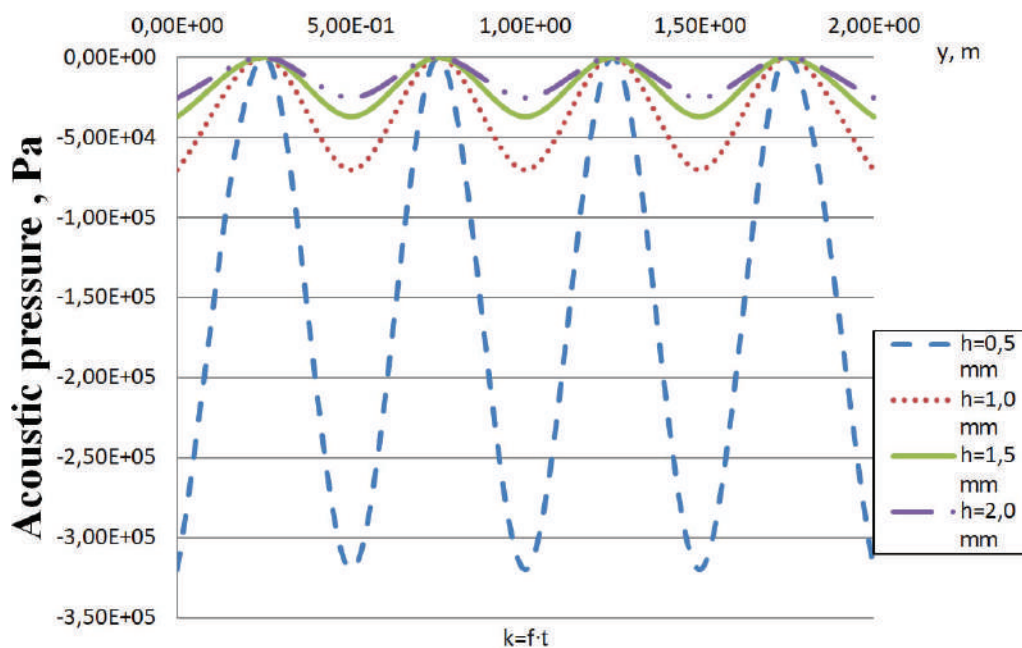


Fig.1. Diagram of the acoustic pressure dependence on the thickness of the plastic layer applied to a metal platform using 3D printing when generating a signal with a single emitter.
 At $B_{-} = 0.3 \text{ T}$, $I_0 = 2 \text{ A}$, $f_i = 0.5 \text{ MHz}$

Fig.1 shows a diagram demonstrating the dependences of acoustic wave formation by a single emitter. There is a sharp decrease in acoustic pressure, with an increase in the value of the layer thickness h .

Analysis of the acoustic vibrations' formation by a common-mode emitter system.

The process of forming acoustic vibrations in the body of the object under test was studied in detail by a number of authors [7, 8, 9, 10, 11, 12]. Based on which, to study the process of forming common-mode wires/emitters of acoustic vibrations by the grid, the effect of the thickness of the layer applied to a metal platform using 3D printing is modelled, with its uniform thickness over the entire

surface of the metal base it is applied on, on the total acoustic pressure created (Fig.2).

Fig.2 shows a diagram demonstrating the formation of an acoustic wave by the grid consisting of emitter wires. Similarly with the diagram shown in Fig.1, there is a sharp decrease in acoustic pressure, with an increase in the value of the layer thickness h .

A study of the effect of the formation of common-mode wires/emitters of acoustic vibrations by the grid was also carried out. The effect of an uneven distribution of the thickness of a layer of plastic applied to a metal platform using 3D printing on the total acoustic pressure created was modelled

If the layer is applied unevenly using 3D printing, the form of the pressure distribution changes dramatically (fig. 3,4)

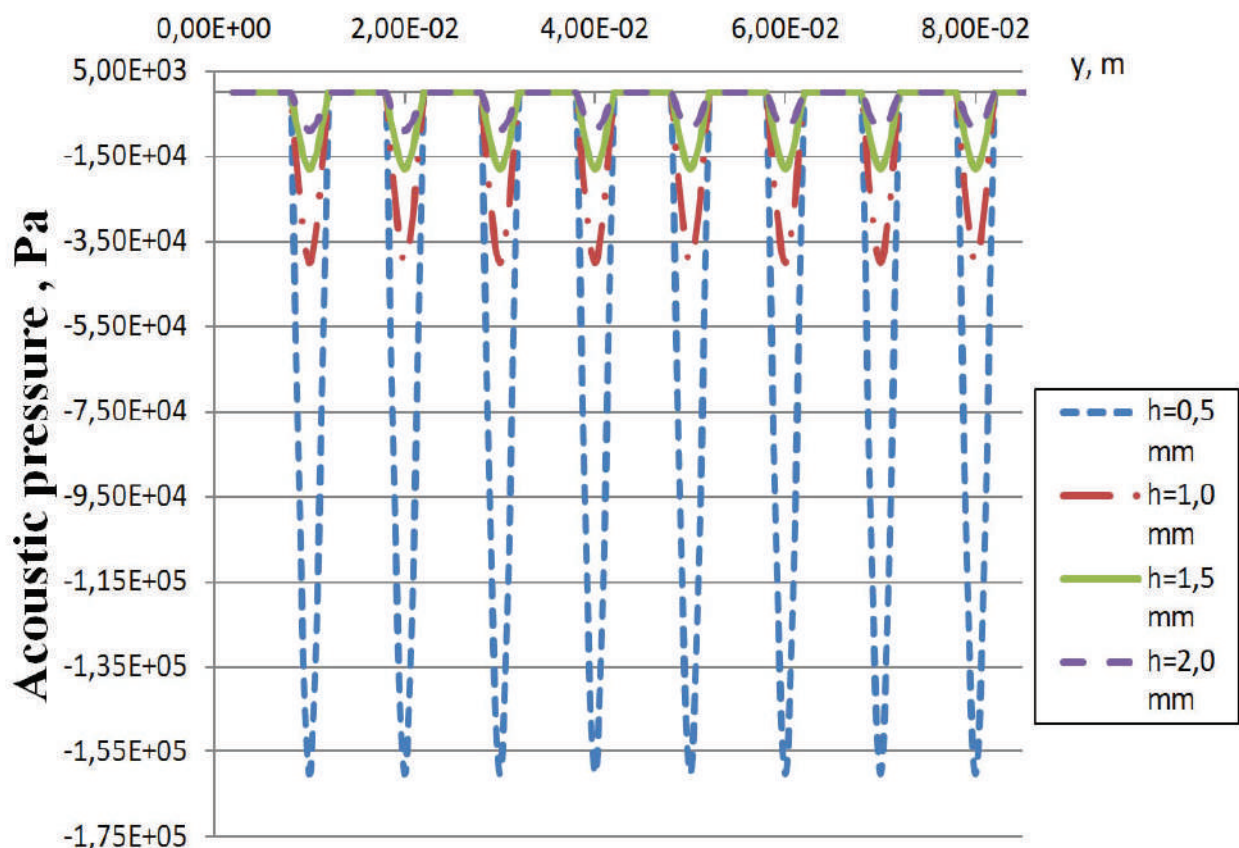


Fig.2. Diagram of the dependence of the total acoustic pressure on the thickness of the plastic layer applied to a metal platform using 3D printing when generating a signal by a family of emitters.

$$AtB_z = 0.37l, I_0 = 2A, f_i = 0.5 \text{ MHz}$$

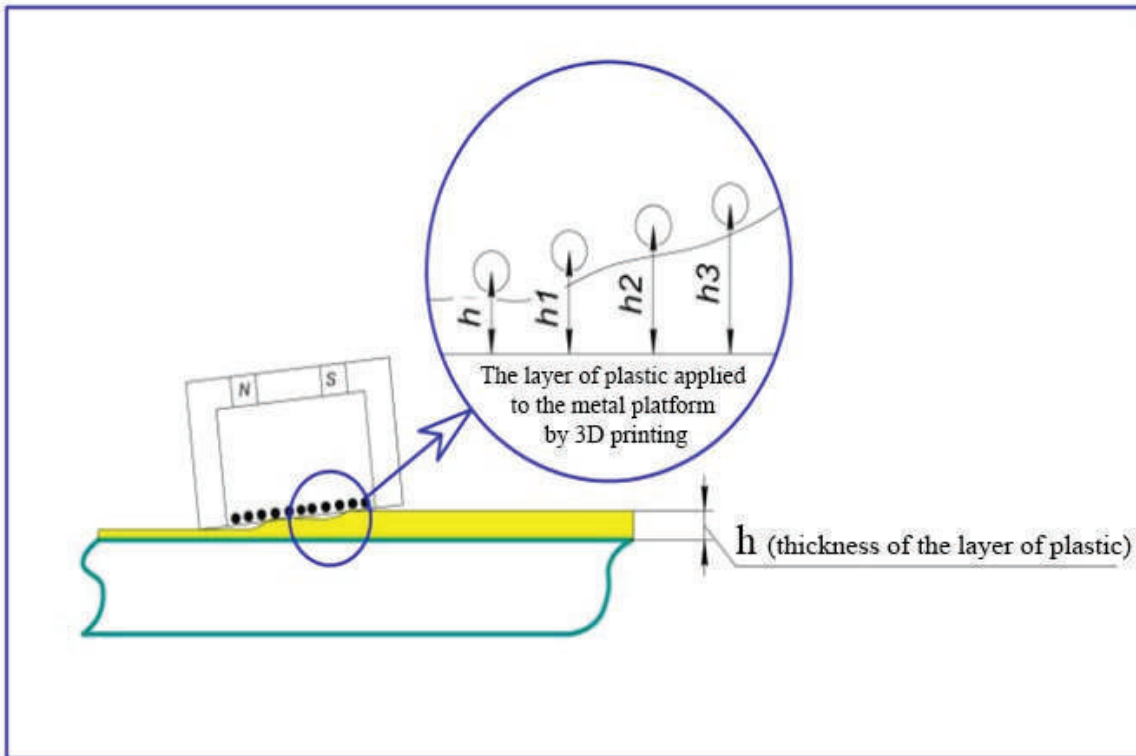


Fig. 3. The EMA method is used to control the thickness of the layer of plastic applied to a metal platform using 3D printing.

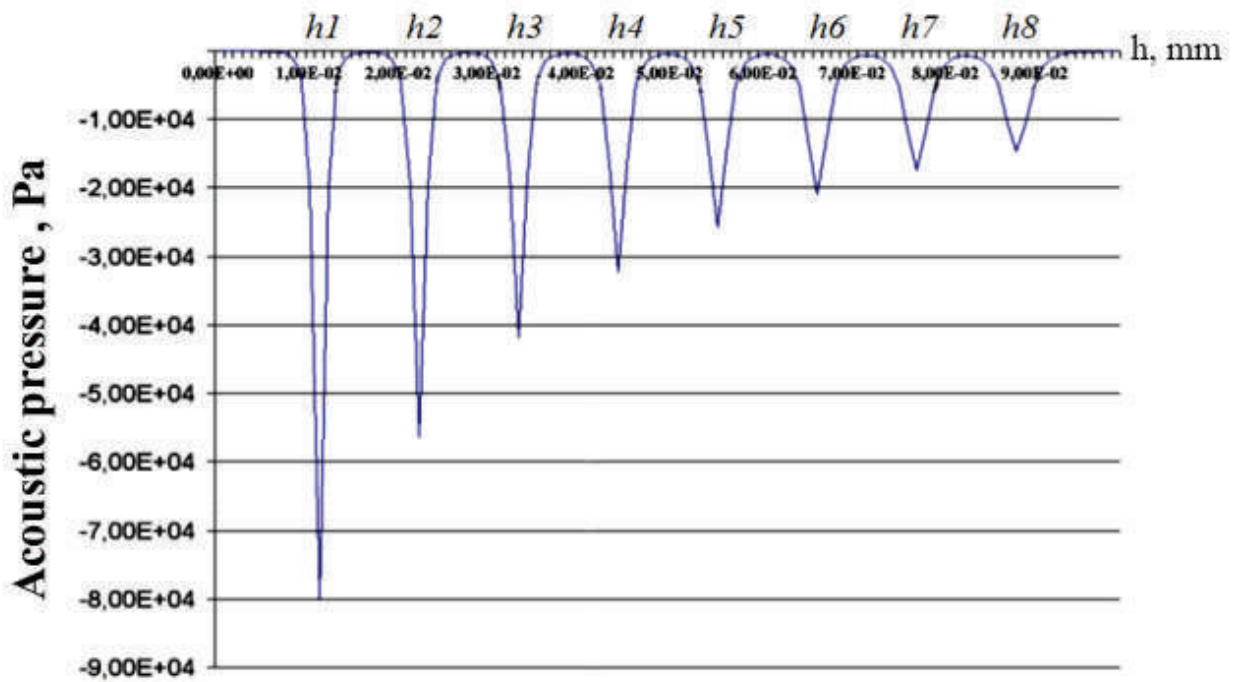


Fig. 4. Dependence of total acoustic pressure at different thickness of plastic layer applied to a metal platform using 3D printing.
 $AtB_z = 0.3TI, I_0 = 2A, f_i = 0.5 \text{ MHz}$

Experimental studies

To confirm the theoretical studies, a number of experiments have been conducted. Universal flaw detector UD4-T and an EMA transducer E411-5-K12 were used during the experiment. The dependence of the acoustic pressure and thickness of the layer applied to the metal platform was analyzed using 3D printing, by applying layers and further measuring the acoustic signal in the area of its application.

As an imitation layer, there were sheets of paper used (with density 80 g/m², thickness 104 microns, or 0.104 mm), which possesses approximate dielec-

tric and magnetic characteristics with plastic used in 3D printing.

The layer's thickness was regulated by the number of layers of paper, not exceeding 1.3 mm. With further increase in thickness, the received signal is disrupted. Two standard 59 and 29 mm thick bars were used as a metal base.

The outcome of the experimental studies is demonstrated in the diagrams of Fig.5 and Fig.6:

The discrepancy between theoretical and practical data does not exceed 5%, which shows good convergence of the outcome of theoretical and experimental studies.

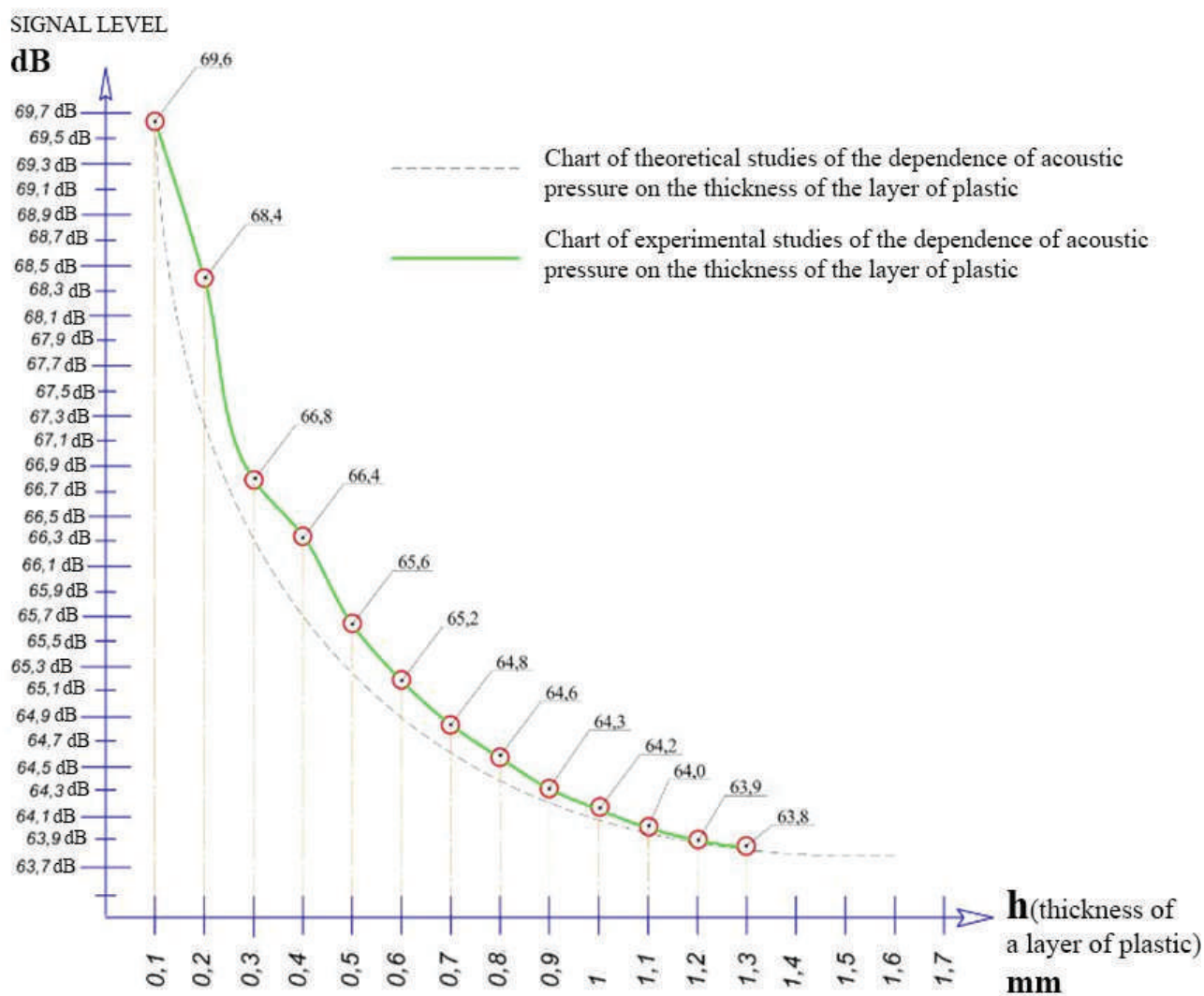


Fig. 5. Diagram of experimental studies of the acoustic pressure's dependence on the thickness of the plastic layer applied to a metal platform using 3D printing, with the 59 mm thick metal base it is applied on.

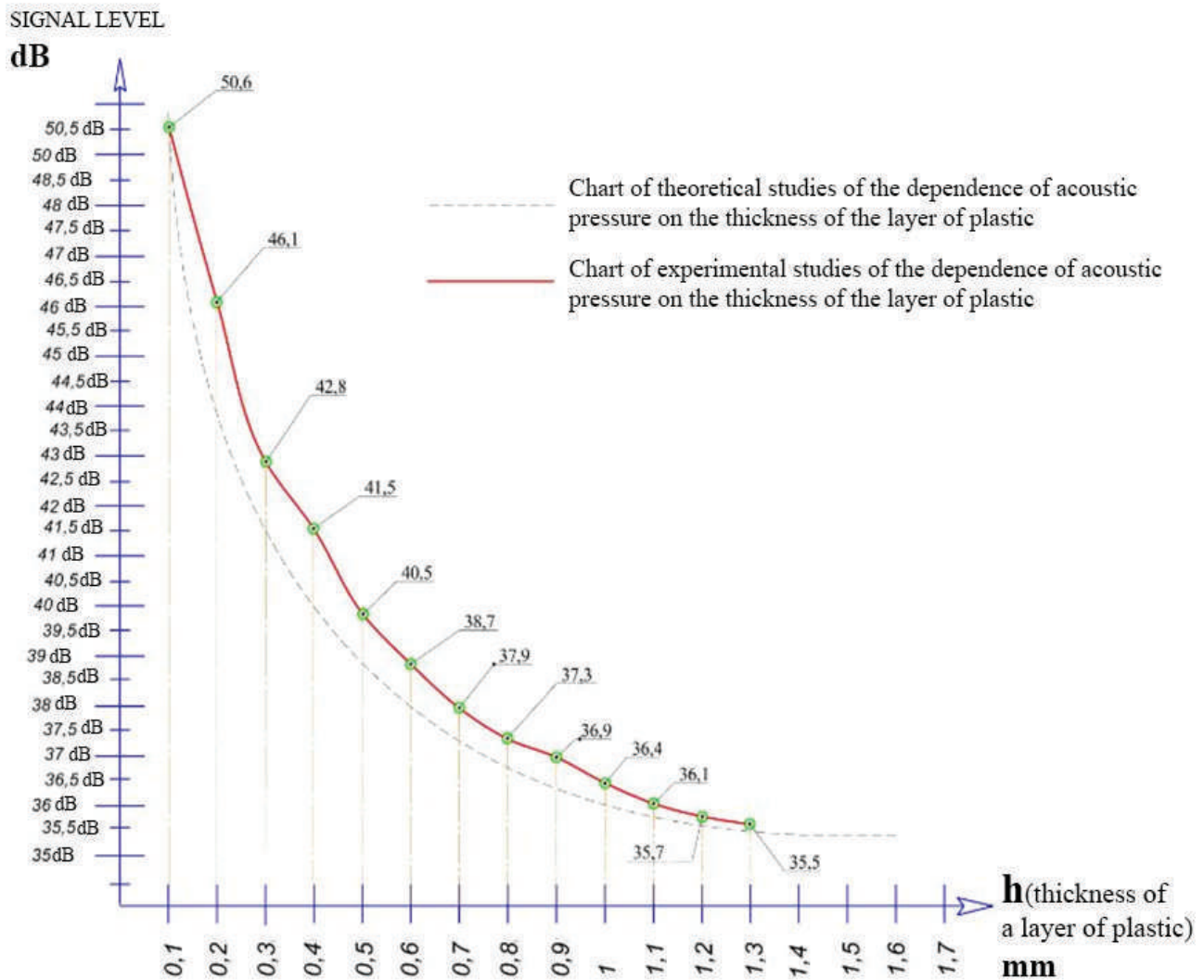


Fig. 6. Diagram of experimental studies of the acoustic pressure's dependence on the thickness of the plastic layer applied to a metal platform using 3D printing, with the 29 mm thick metal base it is applied on.

Summary

The effect of the thickness of the plastic layer applied to a metal platform using 3D printing on the acoustic parameters of the EMA transducer is studied using mathematical modelling.

Dependences showing the effect of the thickness of the plastic layer applied to the metal platform using 3D printing on the created acoustic pressure are obtained experimentally. The maximum acoustic pressure is created when there is no thickness of the plastic layer applied to the metal platform using 3D

printing. The pressure drops sharply, as the layer thickness increases.

Good convergence of the outcome of theoretical and experimental studies is demonstrated, with the approximation error of experimentally obtained data generally not exceeding 5%.

The study is based on a widely proven approach to analyzing the process of acoustic wave formation by an EMA transducer. The reliability of the results obtained is confirmed by the correct use of the mathematical tools and good convergence of the outcome of theoretical and experimental studies.

References

- [1] F. Kliuev et al., Non-Destructive Testing and Diagnostics: Reference Guide, F. Sosnyn et al., eds. Moscow, Russia: Mashinostroenie, 2005, 656 p.
- [2] O. Podolian, "Forming of a magnetic field with given characteristics in EMA converters of systems of nondestructive testing of industrial equipment", Methods and Devices for Quality Control, vol. 17, pp. 18–21, 2006.
- [3] O. Podolian, "EMA converter", RU Patent 2006116939, June 6, 2008.
- [4] H. Tymchyk and O. Podolian, "Formation of special-shaped pulses for electromagnetic acoustic transducers", Bulletin of the NTUU KPI. Ser. Instrument Making, vol. 4, pp. 64–69, 2013.
- [5] B. Yavorskyi and A. Detlaf, Physics Reference. For Engineers and University Students. Moscow, SU: Nauka, 1978.
- [6] Yu. Sazonov, Wave Electromagnetic Acoustic Phenomena in Condensed Media and Physical Methods of Their Usage. Moscow, Russia: GEOS, 2003.
- [7] A. Malynka et al., "Electromagnetic-acoustic method for controlling ferromagnetic sheets and tubes", Defectoscopy, vol. 4, pp. 44–48, 1972.
- [8] V. Komarov, "Modeling of manifestations of electromagnetic acoustic transformation in metals I. Transformation of a quasistationary inhomogeneous electromagnetic field into elastic shear oscillations", Control. Diagnostic, vol. 3, pp. 17–25, 2013.
- [9] V. Tsapenko and Yu. Kuts, The Basis of Ultrasonic Non-Destructive Testing. Kyiv. Ukraine: NTUU KPI, 2010. – 448 p.
- [10] H. Budenkov and O. Korobeinykova, "Influence of chemical composition and metal temperature on the efficiency of electromagnetic acoustic transformation", Defectoscopy, vol. 4, pp. 40–49, 2009.
- [11] O. Podolian et al., "Theoretical Investigations of the Ultrasonic Wave Generation by an Electromagnetic Acoustic Transducer", Kyiv. Ukraine: KPI Science News, 2018, Vol.3(119), pp. 84–92.
- [12] O. Podolian and G. Tymchyk, The Basis of Ultrasonic Non-Destructive Testing. Kyiv. Ukraine: Igor Sikorsky Kyiv Polytechnic Institute, 2020. – 180 p.

Григорій С. Тимчик, Олександр О. Подолян, Максим С. Друзев

КОНТРОЛЬ ТОВЩИНИ ШАРУ ПЛАСТИКУ НАНЕСЕНОГО ЗА ДОПОМОГОЮ ЗД ДРУКУ ЕЛЕКТРОМАГНІТНО-АКУСТИЧНИМ МЕТОДОМ

Проблематика: В даний час в ЕМА метод не застосовується для контролю товщини шару пластику нанесеного на металеву платформу за допомогою ЗД друку, що істотно обмежує області його використання. Разом з тим, широко використовуються ЕМА дефектоскопи, для контролю якості самої металевої платформи. Рішення поставленої задачі дослідження випромінювання дозволить проводити комплексний контроль товщини шару пластику нанесеного за допомогою ЗД друку, що підвищить ефективність ультразвукової дефектоскопії по достовірності і швидкості проведення робіт.

Мета дослідження: У даній статті розглядається можливість визначення товщини шару нанесеного за допомогою ЗД друку використовуючи електромагнітно-акустичний (ЕМА) метод.

Методика реалізації: Проведено аналіз залежності між товщиною шару пластику нанесеного за допомогою ЗД друку на акустичним тиском створюваний в ньому під час контролю з використанням ЕМА перетворювачів. За допомогою математичного моделювання досліджено вплив товщини шару пластику нанесеного на металеву платформу за допомогою ЗД друку на акустичні параметри ЕМА перетворювача. Експериментально отримані залежності, що показують вплив товщини шару нанесеного за допомогою ЗД друку на створюваний акустичний тиск.

Результати дослідження. За допомогою математичного моделювання досліджено вплив товщини шару пластику нанесеного на металеву платформу за допомогою ЗД друку на акустичні параметри ЕМА перетворювача.

Експериментально отримані залежності, що показують вплив товщини шару пластику нанесеного на металеву платформу за допомогою ЗД друку на створюваний акустичний тиск. Максимальний акустичний тиск створюється при відсутності товщини шару пластику нанесеного на металеву платформу за допомогою ЗД друку. Зі збільшенням товщини шару тиск різко падає.

Висновки: Показана хороша збіжність результатів теоретичних і експериментальних досліджень, похибка апроксимації експериментально отриманих даних в основному не перевищує 5%.

В основу досліджень покладений широко апробований підхід до аналізу процесу формування акустичної хвилі ЕМА перетворювачем. Вірогідність отриманих результатів підтверджується коректним використанням математичного апарату й гарною збіжністю результатів теоретичних і експериментальних досліджень.

Подальші дослідження плануються проводити в напрямку дослідження впливу акустичних параметрів електромагнітно-акустичного перетворювача на якість шару пластику нанесеного на металеву платформу за допомогою ЗД друку.

Ключові слова: ЕМА, перетворювач, ЗД, друк, шар, пластик, товщина, акустичний тиск, неруйнівний контроль.

Рекомендована Радою
приладобудівного факультету
КПІ ім. Ігоря Сікорського

Надійшла до редакції
11 липня 2021 року

Прийнята до публікації
14 лютого 2022 року

АВТОРИ НОМЕРА

Барабаш Петро Олексійович
ORCID: 0000-0002-1252-494X

Буштець Яна Миколаївна
ORCID: 0000-0002-3867-7362

Бідюк Петро Іванович
ORCID: 0000-0002-7421-3565

Гао Сінмін
ORCID: 0000-0002-5473-543

Голубєв Олексій Борисович
ORCID: 0000-0002-9629-8352

Друзєв Максим Сергійович
ORCID: 0000-0003-3116-2750

Заболотний Олексій Анатолійович
ORCID: 0000-0002-0913-7789

Завадюк Сергій Вікторович
ORCID: 0000-0002-4338-6244

Куделя Петро Петрович
ORCID: 0000-0003-4941-5795

Кузнєцов Юрій Миколайович
ORCID: 0000-0002-5473-5438

Кузнєцова Наталія Володимирівна
ORCID: 0000-0002-1662-1974

Левенчук Людмила Борисівна
ORCID: 0000-0002-8600-0890

Лобода Петро Іванович
ORCID: 0000-0003-2012-1080

Маслянюк Павло Павлович
ORCID: 0000-0003-4001-7811

Мацевитий Юрій Михайлович
ORCID: 0000-0002-6127-0341

Нехороших Микита Віталійович
ORCID: 0000-0001-6269-0359

Подольян Олександр Олександрович
ORCID: 0000-0003-0522-1403

Притула Наталя Олександрівна
ORCID: 0000-0002-3500-5165

Ремізов Дмитро Олексійович
ORCID: 0000-0002-8330-3562

Романюк Вадим Васильович
ORCID: 0000-0003-3543-3087

Савчук Іван Васильович
ORCID: 0000-0003-1682-1551

Самойленко Олексій Васильович
ORCID: 0000-0002-0403-1707

Середа Володимир Володимирович
ORCID: 0000-0002-4645-3123

Соловійова Тетяна Олександрівна
ORCID: 0000-0003-0222-6890

Соломаха Андрій Сергійович
ORCID: 0000-0002-1061-6277

Степанчук Анатолій Миколайович
ORCID: 0000-0002-0363-5226

Терентьєв Олександр Миколайович
ORCID: 0000-0002-4288-1753

Тимчик Григорій Семенович
ORCID: 0000-0003-1079-998X

Троком Ярослав Євгенович
ORCID: 0000-0001-5672-7717

Трофимчук Олександр Миколайович
ORCID: 0000-0003-3782-4209

Янчевський Ігор Владиславович
ORCID: 0000-0002-7113-2276